PATENT
Customer No. 22,852
Attorney Docket No. 2356.0011-10

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:	)
Marc ALIZON et al.	)
	) Group Art Unit: 1648
Appln. No.: 10/076,370	)
	) Examiner: Parkin, J.
Filed: February 19, 2002	)
For: VARIANT OF LAV VIRUSES	)

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

DECLARATION UNDER 37 C.F.R. § 1.132

I, Dr. Marie-Lise GOUGEON, declare that:

1. I am a Research Director in Pasteur Institute, head of the Antiviral Immunity, Biotherapy and Vaccine Unit.

2. I have read and understood Application Serial No. 10/076,370, including pending claims 23-28 and 31-39, as well as withdrawn claims 40-50, copies of which are attached hereto.

3. I am not an inventor or owner of Application Serial No. 10/076,370, or the subject matter contained therein.

4. I understand that the Examiner has rejected claims 23-28 and 31-39 of this application as allegedly lacking written description support in the specification.

5. I have been asked to comment on the disclosure of the genus of immunogenic polypeptide fragments comprising HIV-1_{MAL} epitopes of 5-150 amino acid

residues, wherein at least one amino acid residue is substituted at one of the specific positions, in Application Serial No. 10/076,370 (the '370 application).

6. Page 23 of the '370 application describes peptides, included in the invention, comprising or consisting of the conserved regions (which have the lengths of 21, 43, 79, 94, and 131, respectively). Describing a peptide spanning positions 680-700 necessarily describes a peptide with a length of 21 amino acid residues.

7. The identity of the mutations encompassed by the disclosure in the '370 application can be ascertained by considering Figure 3. Comparison of the various Env sequences in Figure 3 highlights positions where an amino acid is substituted in all of the sequences designated LAV_{MAL}, ARV2, and LAV_{ELI} when compared to LAV_{BRU}. Amino acid positions substituted in all of the sequences designated LAV_{MAL}, ARV2, and LAV_{ELI} when compared to LAV_{BRU} are called nonconserved amino acids.

8. The fact that the nonconserved amino acids differ in all of the sequences designated LAV_{MAL}, ARV2, and LAV_{ELI} when compared to LAV_{BRU} suggests that each nonconserved amino acid is not required for immunogenicity. Viewing Figure 3, it is clear to me which amino acids are nonconserved. The claims include a listing of these nonconserved amino acids, which are targets for mutation.

9. The '370 application states on page 23 that "[p]roteins containing or consisting of the 'well conserved stretches' are of particular interest for the production of immunogenic compositions and (preferably in relation to the stretches of the <u>env</u> protein) of vaccine compositions against the LAV-1 viruses." This statement implies that nonconserved amino acids, such as those provided in claim 23 as derived from

Figure 3, are not essential for immunogenicity of the peptides, the "well conserved stretches" being of particular interest for the production of immunogenic compositions. This was known for other antigenic systems (Tanabe et al., J. Biochem 96:365 (1984), Lacal et al., Mol Cell Biol 6:1002 (1986)) at the time of filing of this application. Several recent studies support this statement for HIV, both at the cellular (Frahm et al., J. Virol 78:2187 (2004) and humoral (Yang et al., J. Virol. 78:4029 (2004)) level, and the use of consensus sequences in vaccine design to minimize the genetic differences between vaccine strains and contemporary isolates is currently being considered (Gaschen et al., Science 296:2354, 2355, second paragraph (2004); and Gallo, Lancet, early online publication, last paragraph of p.3 (September 28, 2005); Burton et al., Nature Immunol. 5:233-236, 235 'strategies for immunogen design' paragraph (2004))

10. Furthermore, claim 40 recites a structural requirement and a functional requirement for the claimed immunogenic peptides. Claim 40 recites an immunogenic, HIV-1 Env peptide of at least 21 amino acid residues of LAV_{MAL} Env sequences in Figures 3E-3F comprising at least one of the following conserved sequences:

a) amino acid residues 680-700;

b) amino acid residues 488-530;

c) amino acid residues 211-289;

d) amino acid residues 37-130; and

e) amino acid residues 490-620;

wherein said peptide binds to antibodies in AIDS patient sera; and wherein said antibodies are capable of binding to viral antigens encoded by the LAV_{MAL} molecular clone having C.N.C.M. accession number I-641.

11. The structural requirement of the claimed immunogenic peptides in claim 40 is the at least one conserved sequence.

12. The functional requirement of the peptides in claim 40 is that the peptides having the at least one conserved sequence are immunogenic.

13. The specification identifies a correlation between these conserved sequences and this function of the claimed peptides. It states that the peptides containing the well conserved regions are of "particular interest for the production of immunogenic compositions and (preferably in relation to the stretches of the <u>env</u> protein) of vaccine compositions against the LAV-1 viruses." *See* specification, page 23. The specification suggested in 1986 that conserved regions of HIV proteins contribute to the induction of neutralizing antibodies to those proteins. This has been confirmed. See Holmbach et al., J. of Virology 67:1612-1619, 1617, second column, first and last paragraph (1993); Spenlehauer et al., J. of Virology 72:9855-9864, 9862, last paragraph of the first column (1998).

14. I further declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true, and further, that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under

Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Dated: Oct 21, 2005

By: ______________________
Dr. Marie-Lise Gougeon

INSTITUT PASTEUR

Marie-Lise GOUGEON

INSTITUT PASTEUR
Head of the Antiviral Immunity, Biotherapy and Vaccine Unit
Molecular Medicine Department
28 rue du Dr. Roux
75724 Paris Cedex 15

Born PANTHIER
July 15, 1950, Paris
Citizenship France
Married, 3 children

Address: Institut Pasteur
Antiviral Immunity, Biotherapy and Vaccine Unit
Molecular Medicine Department
28 rue du Docteur Roux
75724 PARIS Cedex 15
Tel : 01 45 68 89 07
Fax : 01 45 68 89 09
e-mail : mlgougeo@pasteur.fr

EDUCATION AND DIPLOMA

1972 DES Biochemistry, Paris XI University

1974 : Master's degree in Biochemistry, Paris XI University

1975 : Post-graduate Certification in Biochemistry and Microbiology, Institut Pasteur

1978 : PhD in Immunology . Department of Immunology, Institut Pasteur/PVII University

1986 : Doctorate Thesis in Immunology, Paris VII University

POSITIONS AND AWARDS

1975 :	Roux Fundation Fellowship
1978 :	Research Assistant, Institut Pasteur
1987 :	Chargée de Recherche, l'Institut Pasteur
1987-1990 :	Senior Investigator, Pédiatric Immunology and Rhumatology (INSERM U 132), Hôpital Necker, Paris
1990-93	Co-Director of the Immunology Course/DEA of Pasteur Institute
1993-1995 :	Associate Professor of Immunology, Paris VII University
Since 1996 :	Research Director, Institut Pasteur
Since 2002 :	Head of the « Antiviral Immunity, Biotherapy and Vaccine » Unit, Department of Molecular Medicine, Institut Pasteur

SCIENTIFIC COMMITEES AND EDITORIAL ACTIVITY

- Président of the « European Cell Death Organisation » (ECDO) 1999-2001.
- Member of the coordinated action of ANRS AC21 on T cell homeostasis and HIV (2001-)
- Member of several scientific commitees in ANRS
- Member of the scientific commitee « HIV and therapy », ABBOTT Company (1999-2005)
- Member of the scientific commitee « Lipodystrophy » BMS Company (2000-2004)
- Member of the scientific commitee ATC Biotherapy, INSERM (2002-)
- Associate Editor of *Cell Death and Differentiation* (1997-2003)
- Member of the Editorial Board of *Current Molecular Medicine* (2003-)
- Member of the Editorial Board of *AIDS* (2001-)

CONSULTANT CONTRACTS

- Consultant of Applied Immune Science Company (1993-1995)
- Consultant of ABBOTT Company (1999-2005)
- Consultant of BMS Company (2000-2004)
- Consultant of Bayer Company in 2002
- Expert for the Swedish Research Council (Stockholm) in 2003
- Expert for EU for the 6th PCRDT in 2004

DISTINCTION

1994: • Prize from the French Academy of Science/CEA for the discovery of Programmed Cell Death in AIDS.

PUBLICATIONS IN PEER REVIEW JOURNALS

M. JOSKOWICZ, M.L. GOUGEON, I. LOWY, M. SEMAN and J. THEZE
Helper cells from lymph node or spleen induce different 7S antibody responses.
Ann. Immunol. Inst. Pasteur (1981)132, 97-110.

M.L. GOUGEON,L. LECLERCQ,I. LOWY,G. BISMUTH,G. SOMME and J.THEZE
In vitro inhibition of the helper activity of GAT specific T cell lines by a syngeneic anti-idiotypic serum: preferential effect on the IgG1 response.
Cell. Immunol.* (1982) *71, 254-269.

M. SEMAN, V. ZILBERFARB, M.L. GOUGEON and J. THEZE
Functional analysis of GAT-specific T cell clones : H2 restricted monoclonal T helper cells do not regulate expression of antibody isotypes.
J. Immunol.* (1982)*129, 217-221.

M.L. GOUGEON, G. BISMUTH, L. LECLERCQ, G. SOMME and J. THEZE
Idiotype expression and fine specificity of GAT specific T proliferating cells.
Cell. Immunol.* (1983)*75, 103-110.

M.L. GOUGEON and J. THEZE
MHC-linked Ir gene control of T cell responses : GAT-specific proliferating T cells and helper T cells are elicited in non responder mice immunized with soluble GAT.
J. Immunol.* (1983)*130, 1521-1526.

G. BISMUTH, G. SOMME, C. ROTH, M.L. GOUGEON and J. THEZE
GAT-specific T cells do not express B cell public idiotopes but can be primed by monoclonal anti-idiotypic antibodies.
Eur. J. Immunol.* (1984)*14, 503-510.

1985.1 C. ROTH, G. SOMME, M.L. GOUGEON and J. THEZE
Induction by monoclonal anti-idiotypic antibodies of an anti-poly (Glu^{60} Ala^{30} Tyr^{10}) (GAT) immune response in GAT non responder mice.
Scand. J. Imunol.* (1985)*21, 361-367.

M.L. GOUGEON, G. BISMUTH and J. THEZE
Differential effects of monoclonal antibodies anti L3T4 and anti-LFA-1 on the antigen-induced proliferation of T helper cell clones : correlation between their susceptibility to inhibition and their affinity for antigen.
Cell. Immunol.* (1985)*95, 75-83.

M.L. GOUGEON, L. LECLERCQ, G. BISMUTH and J. THEZE
Differential requirements for the production by T helper clones of the lymphokines involved in the induction of polyclonal proliferation and differentiation of unstimulated B lymphocytes.
J. Immunol.* (1985)*135, 1878-1883.

J. THEZE, L. LECLERCQ, M.L. GOUGEON
T helper cell control of B cell development and isotype expression.
International Reviews of Immunology* (1986)*1, 185-216

M.L. GOUGEON and J. THEZE
A polyclonal assay for T helper function involving T-B cell contact mediated by L3T4 molecules.
J. Immunol.* (1986)*137, 755-759.

M.L. GOUGEON and J. THEZE
Participation of L3T4 in the avidity, activation and interactions of T helper cell clones.

Forum in Immunology, Ann.Inst.Pasteur/16th (1987)***138,144.***

A. DIU, M.L. GOUGEON, J.L. MOREAU, E. REINHERTZ and J. THEZE
Activation of resting human B cells by T helper cell clone supernatant : characterisation of a human B cell activating factor.
Proc. Natl. Acad. Sci. USA* (1987)*84, 9140-9144.

M.L. GOUGEON, G. DREAN, F. LE DEIST, M. DOUSSEAU, M. FEVRIER, A. DIU, J.THEZE, C. GRISCELLI and A. FISCHER
Human Severe Combined Immunodeficiency disease (SCID): Phenotypic and functional characteristics of peripheral blood lymphocytes.
J. Immunol.* (1990)*145, 2873-2879

M. BENKERROU, M.L. GOUGEON, C. GRISCELLI AND A. FISCHER
Hypogammaglobulinémie G et A avec hypergammaglobulinémie M.
Arch. Fr. Pediat.* (1990)*47, 345-349

M.L. GOUGEON, R. OLIVIER, S. GARCIA, D. GUETARD, T. DRAGIC, C DAUGUET, L. MONTAGNIER
Mise en évidence d'un processus d'engagement vers la mort cellulaire par apoptose dans les lymphocytes de patients infectés par le VIH.
C. R. Acad. Sci. Paris,* (1991)*312, 529-537

M.L. GOUGEON, L. MORELET, M. DOUSSAU, J. THEZE, C. GRISCELLI, A. FISCHER
Hyper IgM Immunodeficiency syndrome: Influence of lymphokines on in vitro maturation of peripheral B cells.
J. Clin. Immunol.* (1992)*12, 1-9

M.L. GOUGEON and L. MONTAGNIER
New concepts in the mechanisms of CD4+ lymphocytes depletion in AIDS and the influence of opportunistic infections.
Res. Microbiol.8th Forum in Microbiology* (1992)*362-368

M.L. GOUGEON and L. MONTAGNIER
Programmed Cell Death in T lymphocytes from Human Immunodeficiency virus-infected individuals.
Advances in Allergy and Immunology*, (1992),*1:89-97

M.L. GOUGEON, V. COLIZZI, A. DALGLEISH and L. MONTAGNIER
New concepts in AIDS pathogenesis.
AIDS Res.and Hum. Retrov.* (1993)*9 : 267-269

M.L. GOUGEON, A.G. LAURENT-CRAWFORD, A.G. HOVANESSIAN, L. MONTAGNIER
Direct and indirect mechanisms mediating apoptosis during HIV infection : contribution to in vivo CD4 T cell depletion.
Seminars in Immunology,* (1993) *5 : 306-313

L. MONTAGNIER and M. L. GOUGEON
New concepts in AIDS pathogenesis
***L. Montagnier and M.L. Gougeon Eds, Marcel Dekker* (1993)**

J. HEENEY, R. JONKER, W. KOORNSTRA, R. BUBBES, H. NIPHUIS, A.M. DI RIENZO, M.L. GOUGEON, L.MONTAGNIER (1993)
The resistance of HIV chimpanzees to progression to AIDS correlates with absence of HIV-related T-cell dysfunction.
J. Med. Primatol.* (1993) *22:194-200

M.L. GOUGEON, S. GARCIA, J. HEENEY, R. TSCHOPP, H. LECOEUR, D. GUETARD, V. RAME, C. DAUGUET, L. MONTAGNIER

Programmed Cell Death in AIDS-related HIV and SIV infections.
AIDS Res. and Hum. Retrov.* (1993)*9:553-563

M.L. GOUGEON and L. MONTAGNIER
Apoptosis in AIDS.
Science,* (1993)*260:1269-1270

M.L. GOUGEON, G. DADAGLIO, S. GARCIA, R. ROUE, H. MÜLLER-ALOUF, L. MONTAGNIER
Is a dominant superantigen involved in AIDS pathogenesis?
Lancet,* (1993)*342 : 50-51

M.L. GOUGEON and L. MONTAGNIER.
Programmed cell death and AIDS (Letter)
Science,* (1993). *262: 1355-1357

M. L. GOUGEON and L. MONTAGNIER
Apoptosis in T lymphocytes during HIV infection : influence of superantigens and correlation with AIDS pathogenesis.
Apoptosis II : The molecular basis of cell death.* (1994) *L.D. Tomei and F.O. Cope Eds, Cold Spring Harbor Laboratory Press. pp.5

G. DADAGLIO, S. GARCIA, L. MONTAGNIER, M.L. GOUGEON
Selective anergy of V□8+ T cells in asymptomatic HIV-infected individuals .
J. Exp. Med.* (1994).*179:413-424

M.L. GOUGEON
Does apoptosis contribute to CD4 T cell depletion in HIV infection ?
Cell Death and Differentiation, (1995) 2:1-9

M.L. GOUGEON
Chronic activation of the immune system in HIV infection. Contribution to T cell apoptosis and Vß selective T cell anergy.
Current Topics in Microbiology and Immunology,200:177-193

S. BOULLIER, M. COCHET, F. POCCIA, M.L. GOUGEON
CDR3-independent γδVd1+ T cell expansion in the peripheral blood of HIV-infected persons
J. Immunol. (1995) 154:1418-1431

P.X.PETIT,LECOEUR H, ZORN E.DAUGUET C.MIGNOTTE B.,GOUGEON ML
Alterations in mitochondrial structure and function are.early events of dexamethasone-induced apoptosis.
J. Cell Biology. (1995) 130: 157-167

Y. SASAKI, BLANCHARD A, WATSON HL, GARCIA S, DULIOUST A, MONTAGNIER L, GOUGEON M
In vitro influence of *Mycoplasma penetrans* on activation of peripheral T lymphocytes from healthy donors or HIV-infected individuals.
Infection and Immunity. (1995) 63:4277-4288

J. F. TORRES-ROCCA, LECOEUR H, AMATORE C, GOUGEON ML;
The early intracellular production of a reactive oxygen intermediate mediates apoptosis in dexamethasone-treated thymocytes.
Cell Death and Diff. (1995) 2:309-319

M-L. GOUGEON, LECOEUR H., CALLEBAUT C., JACOTOT E., DULIOUST A., ROUE R.,MONTAGNIER L., HOVANESSIAN A.
Selective loss of CD4+/CD26+ T cell subset during HIV infection;
Res. Immunol. (1996) 147:5-8

F. BOUDET, LECOEUR H, GOUGEON M-L.

Apoptosis associated with ex-vivo down-regulation of Bcl-2 and up-regulation of Fas in potential cytotoxic CD8+ T lymphocytes during HIV infection.
J. Immunol. (1996),156:2282

M-L. GOUGEON, LECOEUR H, DULIOUST A, ENOUF M. G., CROUVOISIER M., GOUJARD C.DEBORD T., MONTAGNIER L.
Programmed cell death in peripheral lymphocytes from HIV-infected persons : the increased susceptibility to apoptosis of CD4 and CD8 T cells correlates with lymphocyte activation and with disease progression;
J. Immunol. (1996), 156:3509

POCCIA F, BOULLIER S, LECOEUR H, COCHER M, COLIZZI V, FOURNIE JJ GOUGEON M-L.

Vg9Vd2 T cell deletion and anergy to non-peptidic mycobacterial antigens in asymptomatic HIV-infected persons.
J. Immunol. (1996), 157:449-461

S. GARCIA, DADAGLIO G., CILOTE V., CHENAL H., BONDURAND A., GOUGEON M
Evidence for an in vivo superantigenic activity in HIV-infected individuals.
Blood (1996), 88:2151-2161

S. GARCIA, DADAGLIO, G., GOUGEON M-L.
Limits of the Human-PBL-SCID Mice model : Severe restriction of the Vß T-cell repertoire of engrafted human T cells.
Blood (1996), 89:329-336

M-L. GOUGEON, LECOEUR H., HEENEY J., BOUDET F.
Comparative analysis of apoptosis in HIV-infecred humans and chimpanzees : relation with lymphocyte activation.
Immunol. Letters (1996) 51:75-81

F. BUSEYNE, FEVRIER M., GARCIA S., GOUGEON M-L, RIVIERE Y.
Dual function of a HIV-specific cytotoxic T lymphocyte clone : inhibition of HIV replication by noncytoloytic mechanisms and lysis of HIV-infected CD4+ cells.
Virology (1996) 225:248-253

H. LECOEUR and M-L. GOUGEON.
Comparative analysis of flow cytometric methods for apoptosis quantitation in thymocytes and human peripheral blood lymphocytes of controls and HIV^{+} persons. Evidence forinterferences of granulocytes and erythrocytes.
J. Immunol. Methods (1996)198:87-99

A. AMENDOLA, GOUGEON M.L., POCCIA F., BONDURAND A., FESUS L. PIACENTINI M.
Induction of tissue transglutaminase in HIV-pathogenesis : evidence for high rate of apoptosis on CD4+ T lymphocytes and accessory cells in lymphoid tissues.
Proc. Natl. Acad. Sci.(USA) (1996) 93:11057

L. MONTAGNIER, BRENNER C, CHAMARET S, GUETARD D, BLANCHARD A, DE SAINT MARTIN J, POVEDA J.D., PIALOUX G., GOUGEON M.L.
HIV infection and AIDS with negative serology.
J. Infect. Disease (1997) 175:955-959

M-L GOUGEON, LECOEUR H., BOUDET F., LEDRU E., MARZABAL S., BOULLIER S., ROUE R., NAGATA S., HEENEY J.
Lack of chronic immune activation in HIV-infected chimpanzees correlates with the resistance of T cells to Fas/Apo-1 (CD95)-induced apoptosis and preservation of a Th1 phenotype.
J. Immunol. (1997) **158:2964**

S. GARCIA, FEVRIER M., DADAGLIO G., LECOEUR H., RIVIERE Y., M-L GOUGEON

Potential deleterious effect of anti-viral cytotoxic lymphocyte through the CD95 (Fas/APO-1)-mediated pathway during chronic HIV infection.
Immunol. Letters (1997) 57:53-58

TUBIANA, R., GOMARD E., FLEURY H., GOUGEON M.L., MOUTHON B., PICOLET H., KATLAMA K.
Vaccine therapy in early HIV-1 infection using a recombinant canarypox virus expressing the gp160 MN (ALVAC HIV) : a double-blind controlled randomized study.
AIDS (1997) 11:819-841

COULAUD J-P, GOUGEON M-L, GOMARD E, DESCAMPS D, LEBON P, ABOULKER J-P, BIZZINI B, ZAGURY D.
A placebo-controlled clinical phase I trial with combined anti-HIV-1 and anti-IFN-a immunization.
AIDS (1997) 11:937-938

Y. POQUET, HALARY F., CHAMPAGNE E., DAVODEAU F., GOUGEON M-L., BONNEVILLE M., FOURNIE J-J.
Human gd T cells in tuberculosis.
Res. Immunol. (1997) 147: 542-549

A. BLANCHARD, MONTAGNIER L. GOUGEON M-L.
Influence of microbial infections on the progression of HIV disease.
Trends in Microbiology (1997), 5: 326-331

S. BOULLIER, DADAGLIO G., LAFEUILLADE A., DEBORD T., GOUGEON M-L.
Vδ1 T cells expanded in the blood throughout HIV infection display a cytotoxic activity and are primed for TNFa and IFNg production but are not selected in lymph nodes.
J. Immunol. (1997) 159:3629-3637

H. LECOEUR, LEDRU E., PREVOST M-C., GOUGEON M-L.
Strategies for phenotyping apoptotic peripheral human lymphocytes comparing ISNT, Annexin-V and 7-AAD cytofluorometric methods.
J. Immunol. Methods (1997) 209:11-20

F. POCCIA, MALKOVSKY M., GOUGEON M-L, BONNEVILLE M., LOPEZ-BOTET M., FOURNIE J-J, COLIZZI V.
γδ T cell activation or anergy during infections : the role of nonpeptidic TCR ligands and HLA class I molecules.
J. Leukoc. Biol. (1997) 62:287-291

.F. POCCIA, CIPRIANI B, VENDETTI S, COLIZZI V, POQUET Y, BATTISTINI L, LOPEZ-BOTET M, FOURNIE J-J, GOUGEON M-L.
CD94/NKG2 inhibitory receptor complex modulates both anti-viral and anti-tumoral responses of polyclonal phosphoantigen-reactive Vg9Vd2 T lymphocytes.
J. Immunology, 159:6009-6017, 1997.

E. LEDRU, LECOEUR H, GARCIA S, DEBORD T, GOUGEON M-L.
Differential susceptibility to activation-induced apoptosis among peripheral Th1subsets. Correlation with Bcl-2 expression and consequences for AIDS pathogenesis.
J. Immunology (1998), 160: 3194-3206

F. POCCIA, GOUGEON M-L, BONNEVILLE M., LOPEZ-BOTET M, MORETTA A, BATTISTINI L, WALLACE M, COLIZZI, V MALKOVSKY.
Innate T-cell immunity to nonpeptidic antigens.
Immunol. Today (1998), 19:253-256

L. WEISS, ROUX A., GARCIA S., DEMOUCHY C., HAEFFNER-CAVAILLON N, KAZATCHKINE M.D., GOUGEON M-L.

Persistent expansion of V□-restricted double positive CD4$^+$CD8 T lymphocytes in an HIV-infected individual, that express cytotoxicity-associated molecules and are committed to IFN□ and TNF□ production.
J. Inf. Disease (1998), 178:1158-1162

H. LECOEUR, LEDRU E, GOUGEON M-L.
A cytofluorometric method for the simultaneous detection of both intracellular and surface antigens on apoptotic peripheral lymphocytes.
J. Immunol. Methods (1998), 217:11-26

Z. SZONDY, H. LECOEUR, L. FESUS, M-L GOUGEON.
All trans retinoic acid inhibition of anti-CD3 induced T cell apoptosis in HIV infection mostly concerns CD4 T lymphocytes and is mediated via regulation of CD95L expression.
J. Inf. Disease (1998), 178:1288-1298

S. BOULLIER, Y. POQUET, F; HALARY, M. BONNEVILLE, J-J FOURNIE, GOUGEON M-L.
Phosphoantigen activation induces surface translocation of intracellular CD94/NKG2A class I receptor on CD94$^-$ peripheral Vγ9Vδ2 T cells but not on CD94$^-$ thymic or mature □□ T cell clones.
Eur. J. Immunol.(1998), 28:3339-3410

O. TERRADIMMOS, T. POLLICINO, H. LECOEUR, M. TRIPODI, M-L GOUGEON, P. TIOLLAIS, M-A BUENDIA.
p53-independent apoptotic effects of the hepatitis B virus Hbx protein in vivo and in vitro.
Oncogene (1998), 17:2115-2123

S. BOULLIER, Y. POQUET, T. DEBORD, J-J FOURNIE, GOUGEON M-L.
Regulation by cytokines (IL-12, IL-15, IL-4 and IL-10) of theVg9Vd2 response to mycobacterial phosphantigens in responder and anergic HIV-infected persons.
Eur. J. Immunol. (1999) 29:90-99

M-L GOUGEON, S. BOULLIER, V. COLIZZI, F. POCCIA.
Control by Natural Killer Receptor (NKR) of γδT cell immunity to virus.
Microbes and Infection (special issue on γδ T cells) (1999),1:219-226

M-L GOUGEON, H. LECOEUR, Y. SASAKI.
Apoptosis and the CD95 system in HIV disease. Impact of HAART.
Immunol. Lett. (1999), 66:97-103

H. NAORA, M-L GOUGEON
Activation, survival and apoptosis of CD45R0+ and CD45R0- T cells of human immunodeficiency virus-infected individuals: effects of IL-15 and comparison with IL-2.
Immunology, (1999), 97:181-187

H. NAORA, M-L GOUGEON
Enhanced survival and potent expansion of the natural killer cell population of HIV-infected individuals by exogenous IL-15.
Immunol. Lett. (1999), 68:359-367

F POCCIA, L BATTISTINI, B CIPRIANI, G MANCINO, F MARTINI, M-L GOUGEON, V COLIZZI
Phosphoantigen-reactive Vγ9Vδ2 T lymphocytes suppress in vitro HIV-1 replication by cell released anti-viral factors including c-c-chemokines.
J. Inf. Disease (1999), 180: 180:858-61

N.CHRISTEFF, J-C MELCHIOR, E. LEDRU, O. PATEY, P. de TRUCHIS, C. PERONNE, E. A. NUNEZ and M-L GOUGEON·
Lipodystrophy defined by a clinical score in HIV-infected men on HAART: correlation between dyslipidaemia and streroid hormone alterations.
AIDS (1999),13:2251-60

H NAORA, M-L GOUGEON`

IL-15 is a potent survival factor in the prevention of spontaneous but not CD95-induced apoptosis in CD4 and CD8 T lymphocytes of HIV-infected individuals. Correlation with its ability to increase Bcl-2 expression.
Cell Death and Differentiation (1999), 6:1002-11

M-L GOUGEON, L MONTAGNIER

Programmed cell death as a mechanism of CD4 and CD8 T cell deletion in AIDS. Molecular control and effect of highly active anti-retroviral therapy
Annals of the New York Academy of Sciences, (1999), 887:199-212

L. BROWN, SOUBERBIELLE B.E., WESTBY M., MARRIOTT J.B., DESSELBERGER, U. KAY T., GOUGEON M-L, DALGLEISH A.G.

The conserved carboxy terminal region of HIV-1 gp120 is recognized by seron negative HIV-exposed people.
AIDS (1999),13:2515-21

P.MICHEL, A TOURE BALDE, C ROUSSILHON, G ARIBOT, J-L SARTHOU, M-L GOUGEON

Reduced immune activation and T cell apoptosis in HIV-2 compared to HIV-1 infected West African patients. Correlation of T cell apoptosis with ß2 microglobulin concentration and disease evolution.
J. Inf. Diseases, (2000), 181:64-75

E LEDRU, N CHRISTEFF, O PATEY, J-C MELCHIOR, M-L GOUGEON

Alteration of TNF T cell homeostasis following HAART. Implication in the development of HIV-associated lipodystrophy syndrome.
Blood (2000) 95:3191-98

M-L GOUGEON, F POCCIA, S BOULLIER

Human T lymphocytes in HIV disease: effector functions and control by natural killer receptors
Springer Seminars in Immunopathology (2000) 22 :251

M MALKOVSKY, M WALLACE, JJ FOURNIE, P FISCH, F POCCIA, ML GOUGEON

Alternative cytotoxic effector mechanisms in infections with immunodeficiency viruses : γδ T lymphocytes and natural killer cells
AIDS (2000), 14 :S175

ML GOUGEON, E LEDRU, H NAORA, M BOCCHINO, H LECOEUR

HIV, cytokines and programmed cell death : a subtle interplay.
Annals of the New York Academy of Sciences (2000), 926 :30

M CHAMI, D GOZUACIK, K SAIGO, T CAPIOD, P FALSON, H LECOEUR, T URASHIMA. J BECKMAN, M-L GOUGEON, M CLARET, M LE MAIRE, C BRECHOT, P PATERLINI-BRECHOT

Hepatitis B virus-related insertional mutagenesis implicates SERCA1 gene in the control of apoptosis.
Oncogene, (2000), 19 :2877

N CHRISTEFF. EA NUNEZ, M-L GOUGEON. Changes in cortisol/DHEA ratio in HIV-infected men are related to immunological and metabolic perturbations leading to malnutrition and lipodystrophy. [
Annals of the New York Academy of Sciences (2000) 917:962-70,

H LECOEUR, MC PRÉVOST AND M-L GOUGEON

Oncosis is associated to exposure of phosphatidylserine residues on the outside layer of plasma membrane. A reconsideration of the specificity of the AnnexinV-Propidium Iodide assay
Cytometry (2001) 44 :65

M BOCCHINO, E LEDRU, T DEBORD, M-L GOUGEON.

Increased priming for IL-12 and TNF-□ in CD64 monocytes in HIV-1 infection : Modulation by cytokines and therapy.
AIDS (2001) 15 :1213-23

M CHAMI, GOZUACIK D, LAGORCE D,. BRINI M, FALSON P, PEAUCELLIER G, PINTON P, LECOEUR H, GOUGEON ML, LE MAIRE M, RIZZUTO R, BRECHOT C, PATERLINI-BRECHOT P.
SERCA1 truncated proteins unable to pump calcium reduce the endoplasmic reticulum calcium concentration and induce apoptosis.
Journal of Cell Biology (2001) 153:1301-1313

H LECOEUR, M FEVRIER, S GARCIA, Y RIVIERE AND M-L GOUGEON
A novel flow cytometric assay for quantitation and multiparametric characterization of cell-mediated cytotoxicity.
J. Immunological Methods (2001) 253 : 177-87

ML GOUGEON, ROUZIOUX C, LIBERMAN I, BURGARD M. VIARD J-P, BOUCHENAFA K, CAPITANT C, DELFRAISSY JF, THE ANRS 048 GROUP, LEVY Y.
Immunological and virological effects of long-term IL-2 therapy in HIV-1 infected patients.
AIDS,(2001) 15 :1729-32

FELLAY B. CHOFFLON M. JUILLARD C. PAUNIER AM. LANDIS T. ROTH S. GOUGEON ML.
Beneficial effect of co-polymer 1 on cytokine production by CD4 T cells in multiple sclerosis.
Immunology (2001) 104(4):383-391

N CHRISTEFF, JC MELCHIOR, P de TRUCHIS, C PERRONNE, M-L GOUGEON
Increased serum interferon alpha and cortisol : DHEA ratio in antiretroviral-treated HIV-infected men with lipodystrophy are associated with hyperlipidaemia
Eur. J. Clinical Investigation (2002), 32 :43-50

TERRADILLOS O. DE LA COSTE A. POLLICINO T. NEUVEUT C. SITTERLIN D. LECOEUR H. GOUGEON ML. KAHN A. BUENDIA MA.
The hepatitis B virus X protein abrogates Bcl-2-mediated protection against Fas apoptosis in the liver.
Oncogene (2002) 21(3):377-386

LM DE OLIVEIRA PINTO, S GARCIA, H LECOEUR, C RAPP and ML GOUGEON
Increased sensitivity of T lymphocytes to TNFR1 and TNFR2-mediated apoptosis in HIV-infection. Relation to expression of Bcl-2 and active caspase-8 and caspase-3.
Blood, (2002) 99 :1666-1675

ML GOUGEON, M MALKOVSKY, R CASETTI, C AGRATI, F POCCIA
Innate T cell immunity to HIV infection. Immunotherapy with phosphocarbohydrates, a novel strategy of immune intervention?
Vaccine (2002) 20 :1938-1941

LM DE OLIVEIRA PINTO, H LECOEUR, E LEDRU, C RAPP, O PATEY and ML GOUGEON
Lack of control of T cell apoptosis under HAART. Influence of therapy régiment in vivo and in vitro.
AIDS (2002) 16 :329-339

GOUGEON ML, LECOEUR H
Special JIM Issue on Evaluation of Apoptosis
J. Immunol. Meth.(2002), vol. 265

LECOEUR H , DE OLIVEIRA PINTO L, GOUGEON ML.
Multiparameteric flow cytometric analysis of biochemal and functional events associated with apoptosis and oncosis using the 7-AAD.
J. Immunol. Meth.(2002), 265 :81-96

CHRISTEFF N, DE TRUCHIS P, MELCHIOR JC, PERRONNE C, GOUGEON ML.
Longitudinal evolution of HIV-1 associated lipodystrophy is correlated to serum cortisol :DHEA ratio and IFN□
Eur. J. Clin. Invest. (2002) 32(1):43-50

POCCIA F. GOUGEON ML. AGRATI C. MONTESANO C. MARTINI F. PAUZA CD. FISCH P. WALLACE M. MALKOVSKY M.
Innate T-cell immunity in HIV infection: the role of Vgamma9Vdelta2 T lymphocytes.
Current Molecular Medicine.(2002) 2(8):769-81

LEDRU E., FEVRIER M., LECOEUR H., GARCIA S., BOULLIER S., GOUGEON ML
A nonsecreted variant of IL-4 is associated with apoptosis : implication for the T helper-2 polarization in HIV infection.
Blood (2003) 101 : 3102-3105

GOUGEON ML, KROEMER G.
Charming to death : caspase-dependent or-independent ?
Cell Death Diff. (2003), 1-3

GOUGEON ML
Apoptosis as an HIV strategy to escape immune attack
Nature Rev. Immunol.(2003) 3 :392-405

GOUGEON ML, PENICAUD L, FROMENTY B, LECLERCQ P, VIARD JP, CAPEAU J.
Adipocytes targets and actors in the pathogenesis of HIV-associated lipodystrophy and metabolic alterations.
Antivir Ther. 2004 Apr;9(2):161-77

GOUGEON ML
Apoptotic pathways triggered by HIV and consequences on T cell homeostasis and HIV-specific immunity.
Prog Mol Subcell Biol. 2004;36:95-115

GOUGEON ML.
To kill or be killed: how HIV exhausts the immune system.
Cell Death Differ. 2005 S1 :845-54

V. VIGNARD, B. LEMERCIER, A. LIM, MC PANDOLFINO, Y. GUILLOUX, A. KHAMMARI, C.RABU, K. ECHASSERIEAU, F. LANG, ML GOUGEON, B. DRENO, F. JOTEREAU AND N.LABARRIERE.
In vivo Melan-A repertoire expansion following adoptive transfer of highly tumor-reactive Melan-A specific CTL clones in melanoma patients1.
The Journal of Immunology, oct 2005

C LAGRESLE-PEYROU, F YATES, M MALASSIS-SERIS, C HUE, E MORILLON, A GARRIGUE, A LIU, P HAJDARI, D STOCKOLM, O DANOS, B LEMERCIER, M-L GOUGEON, F RIEUX-LAUCAT, J-P dE VILLARTAY, A FISCHER, M CAVAZZANA-CALVO.
Long Term Immune Reconstitution In Rag-1 Deficient Mice Treated By Retroviral Gene Therapy : A Balance between Efficiency and Toxicity.
Blood, Sept 2005

1-22. (Canceled)

23. (Previously Presented) An immunogenic, HIV-1 Env peptide of 5-150 amino acid residues of LAV_{MAL} Env sequences in Figures 3E-3F having at least one amino acid substitution at one or more of positions 8, 9, 90, 102, 131, 133, 140, 156, 172, 177, 179, 185, 188, 192, 198, 207, 209, 290, 305, 308, 323, 333, 335, 337, 341, 342, 353, 356, 359, 363, 404, 428, 440, 457, 41, 477, 483, 484, 486, 538, 555, 641, 652, 656, 660, 663, 694, 740, 733, 799, 854, 856, 862, and 875 in relation to Env sequences of LAV_{BRU}, LAV_{ARV2}, and LAV_{ELI}; wherein said peptide binds to antibodies in AIDS patient sera; and wherein said antibodies are capable of binding to viral antigens encoded by the LAV_{MAL} molecular clone having C.N.C.M. accession number I-641.

24. (Previously Presented) The peptide of claim 23, wherein said peptide is generated by chemical cleavage.

25. (Previously Presented) The peptide of claim 23, wherein said peptide is expressed from a recombinant DNA.

26. (Previously Presented) The peptide of claim 23, wherein said peptide is generated by chemical synthesis.

27. (Previously Presented) A method for detecting antibodies to HIV in a test sample comprising:

a) providing at least one peptide of claim 23 affixed to a solid support;

b) combining a test sample with the at least one peptide affixed to the solid support;

c) optionally rinsing the solid support to remove unbound antibodies of the test sample; and

d) detecting peptide-antibody complex formed, which is indicative of the presence of HIV antibodies in the test sample.

28. (Previously Presented) A method of eliciting neutralizing antibodies to HIV in a mammal comprising:

a) providing a composition comprising at least one peptide of claim 23, a suitable pharmaceutically or physiologically acceptable carrier, and optionally an adjuvant;

b) immunizing the mammal with the composition; and

c) optionally testing a blood sample from the mammal to assay for the binding affinity and neutralizing activity of the elicited antibodies.

29-30. (Canceled)

31. (Previously Presented) The peptide of claim 23, wherein the peptide comprises at least one of the following conserved sequences: positions 37-130, 211-289, 488-530, 490-620, and 680-700 of Env as shown in Fig. 3E-F.

32. (Previously Presented) A method for detecting antibodies to HIV in a test sample comprising:

a) providing at least one peptide of claim 31 affixed to a solid support;

b) combining a test sample with the at least one peptide affixed to the solid support;

c) optionally rinsing the solid support to remove unbound antibodies of the test sample; and

d) detecting peptide-antibody complex formed, which is indicative of the presence of HIV antibodies in the test sample.

33. (Previously Presented) A method of eliciting neutralizing antibodies to HIV in a mammal comprising:

a) preparing a vaccine comprising at least one peptide of claim 31, a suitable pharmaceutically or physiologically acceptable carrier, and optionally an adjuvant;

b) immunizing the mammal with the vaccine; and

c) optionally testing a blood sample from the mammal to assay for the binding affinity and neutralizing activity of the elicited antibodies.

34. (Previously Presented) An immunogenic, HIV-1 Env peptide of at least 21 amino acid residues of LAV_{MAL} Env sequences in Figures 3E-3F having at least one amino acid substitution consisting of an amino acid substitution at one or more of positions 8, 9, 90, 102, 131, 133, 140, 156, 172, 177, 179, 185, 188, 192, 198, 207, 209, 290, 305, 308, 323, 333, 335, 337, 341, 342, 353, 356, 359, 363, 404, 428, 440, 457, 41, 477, 483, 484, 486, 538, and 555, 641, 652, 656, 660, 663, 694, 740, 733, 799, 854, 856, 862, 875; wherein said peptide binds to antibodies in AIDS patient sera; and wherein said antibodies are capable of binding to viral antigens encoded by the LAV_{MAL} molecular clone having C.N.C.M. accession number I-641.

35. (Previously Presented) The peptide of claim 34, wherein the peptide has 21 amino acids.

36. (Previously Presented) The peptide of claim 34, wherein the peptide has 43 amino acids.

37. (Previously Presented) The peptide of claim 34, wherein the peptide has 79 amino acids.

38. (Previously Presented) The peptide of claim 34, wherein the peptide has 94 amino acids.

39. (Previously Presented) The peptide of claim 34, wherein the peptide has 131 amino acids.

40. (Withdrawn) An immunogenic, HIV-1 Env peptide of at least 21 amino acid residues of LAV_{MAL} Env sequences in Figures 3E-3F comprising at least one of the following conserved sequences:

a) amino acid residues 680-700;

b) amino acid residues 488-530;

c) amino acid residues 211-289;

d) amino acid residues 37-130; and

e) amino acid residues 490-620;

wherein said peptide binds to antibodies in AIDS patient sera; and wherein said antibodies are capable of binding to viral antigens encoded by the LAV_{MAL} molecular clone having C.N.C.M. accession number I-641.

41. (Withdrawn) A method for detecting antibodies to HIV in a test sample comprising:

a) providing at least one isolated HIV-1 LAV_{MAL} Env peptide consisting of 5-150 amino acid residues as set forth in Figures 3E-3F affixed to a solid support, wherein said peptide contains a LAV_{MAL}-specific antigenic determinant;

b) combining a test sample with the at least one peptide affixed to the solid support;

c) optionally rinsing to remove unbound antibodies of the test sample;

d) detecting peptide-antibody complex formed, which is indicative of the presence of HIV antibodies in the test sample.

42. (Withdrawn) The method of claim 41, wherein said peptide is generated by chemical cleavage.

43. (Withdrawn) The method of claim 41, wherein said peptide is expressed from a recombinant DNA.

44. (Withdrawn) The method of claim 41, wherein said peptide is generated by chemical synthesis.

45. (Withdrawn) The method of claim 41, wherein said peptide binds to antibodies in AIDS patient sera; and wherein said antibodies are capable of binding to viral antigens encoded by the LAV_{MAL} molecular clone having C.N.C.M. accession number I-641.

46. (Withdrawn) A method of eliciting neutralizing antibodies to HIV in a mammal comprising:

a) providing a vaccine comprising the an isolated HIV-1 LAV_{MAL} Env peptide consisting of 5-150 amino acid residues as set forth in Figures 3E-3F, wherein said peptide contains a LAV_{MAL}-specific antigenic determinant, at least one suitable buffer, and optionally an adjuvant;

b) immunizing the mammal with the vaccine;

c) optionally testing a blood sample from the mammal to assay for the affinity and activity of the antibodies.

47. (Withdrawn) The method of claim 46, wherein said peptide is generated by chemical cleavage.

48. (Withdrawn) The method of claim 46, wherein said peptide is expressed from a recombinant DNA.

49. (Withdrawn) The method of claim 46, wherein said peptide is generated by chemical synthesis.

50. (Withdrawn) The method of claim 46, wherein said peptide binds to antibodies in AIDS patient sera; and wherein said antibodies are capable of binding to viral antigens encoded by the LAV_{MAL} molecular clone having C.N.C.M. accession number I-641.

A VARIANT OF LAV VIRUSES

BACKGROUND OF THE INVENTION

The present invention relates to a virus capable of inducing lymphadenopathies (hereinafter "LAS") and acquired immuno-depressive syndromes (hereinafter "AIDS"), to antigens of this virus, particularly in a purified form, and to a process for producing these antigens, particularly antigens of the envelope of this virus. The invention also relates to polypeptides, whether glycosylated or not, produced by the virus and to DNA sequences which code for such polypeptides. The invention further relates to cloned DNA sequences hybridizable to genomic RNA and DNA of the lymphadenopathy associated virus (hereinafter "LAV") of this invention and to processes for their preparation and their use. The invention still further relates to a stable probe including a DNA sequence which can be used for the detection of the LAV virus of this invention or related viruses or DNA proviruses in any medium, particularly biological, and in samples containing any of them.

An important genetic polymorphism has been recognized for the human retrovirus which is the cause of AIDS and other diseases like LAS, AIDS-related complex (hereinafter "ARC") and probably some encephalopathies (for review, see Weiss, 1984). Indeed all of the isolates, analyzed until now, have had distinct restriction maps, even those recovered at the same place and time [Benn et al., 1985]. Identical restriction maps have only been observed for the first two isolates which were designated LAV [Alizon et al., 1984] and human T-cell lymphotropic virus type 3 (hereinafter "HTLV-3") [Hahn et al., 1984] and which appear to be exceptions. The

genetic polymorphism of the AIDS virus was better assessed after the determination of the complete nucleotide sequence of LAV [Wain-Hobson et al., 1985], HTLV-3 [Ratner et al., 1985 ; Muesing et al., 1985] and a third isolate designated AIDS-associated retrovirus (hereinafter "ARV2") [Sanchez-Pescador et al., 1985]. In particular, it appeared that, besides the nucleic acid variations responsible for the restriction map polymorphism, isolates could differ significantly at the protein level, especially in the envelope (up to 13 % of difference between ARV and LAV), by both amino acids substitutions and reciprocal insertions-deletions [Rabson and Martin, 1985].

Nevertheless, such differences did not go so far as to destroy the immunological similarity of such isolates as evidenced by the capabilities of their similar proteins, (e.g., core proteins of similar nature, such as the p25 proteins, or similar envelope glycoproteins, such as the 110-120 kD glycoproteins) to immunologically cross-react. Accordingly, the proteins of any of said LAV viruses can be used for the _in vitro_ detection of antibodies induced _in vivo_ and present in biological fluids obtained from individuals infected with the other LAV variants. Therefore, these viruses are grouped together as a class of LAV viruses (hereinafter "LAV-1 viruses").

SUMMARY OF THE INVENTION

In accordance with this invention, a new virus has been discovered that is responsible for diseases clinically related to AIDS and that can be classified as a LAV-1 virus but that differs genetically from known LAV-1 viruses to a much larger extent than the known LAV-1 viruses differ from each other. The new virus

is basically characterized by the cDNA sequence which is shown in Figures 7A to 7I, and this new virus is hereinafter generally referred to as "LAV_{MAL}".

Also in accordance with this invention, variants of the new virus are provided. The RNAs of these variants and the related cDNAs derived from said RNAs are hybridizable to corresponding parts of the cDNA of LAV_{MAL}. The DNA of the new virus also is provided, as well as DNA fragments derived therefrom hybridizable with the genomic RNA of LAV_{MAL}, such DNA and DNA fragments particularly consisting of the cDNA or cDNA fragments of LAV_{MAL} or of recombinant DNAs containing such cDNA or cDNA fragments.

DNA recombinants containing the DNA or DNA fragments of LAV_{MAL} or its variants are also provided. It is of course understood that fragments which would include some deletions or mutations which would not substantially alter their capability of also hybridizing with the retroviral genome of LAV_{MAL} are to be considered as forming obvious equivalents of the DNA or DNA fragments referred to hereinabove.

Cloned probes are further provided which can be made starting from any DNA fragment according to the invention, as are recombinant DNAs containing such fragments, particularly any plasmids amplifiable in procaryotic or eucaryotic cells and carrying said fragments. Using cloned DNA containing a DNA fragment of LAV_{MAL} as a molecular hybridization probe - either by marking with radionucleotides or with fluorescent reagents - LAV virion RNA may be detected directly, for example, in blood, body fluids and blood products (e.g., in antihemophylic factors such as Factor VIII concentrates). A suitable method for achieving such detection comprises immobilizing LAV_{MAL} on a support (e.g., a nitrocellulose filter), disrupting the virion and

hybridizing with a labelled (radiolabelled or "cold" fluorescent- or enzyme-labelled) probe. Such an approach has already been developed for Hepatitis B virus in peripheral blood (according to Scotto J. et al. Hepatology (1983), 3, 379-384).

Probes according to the invention can also be used for rapid screening of genomic DNA derived from the tissue of patients with LAV related symptoms to see if the proviral DNA or RNA present in their tissues is related to LAV_{MAL}. A method which can be used for such screening comprises the following steps : extraction of DNA from tissue, restriction enzyme cleavage of said DNA, electrophoresis of the fragments and Southern blotting of genomic DNA from tissues and subsequent hybridization with labelled cloned LAV provival DNA. Hybridization *in situ* can also be used. Lymphatic fluids and tissues and other non-lymphatic tissues of humans, primates and other mammalian species can also be screened to see if other evolutionary related retroviruses exist. The methods referred to hereinabove can be used, although hybridization and washings would be done under non-stringent conditions.

The DNA according to the invention can be used also for achieving the expression of LAV viral antigens for diagnostic purposes, as well as for the production of a vaccine against LAV. Fragments of particular advantage in that respect will be discussed later. The methods which can be used are multifold :

a) DNA can be transfected into mammalian cells with appropriate selection markers by a variety of techniques, such as calcium phosphate precipitation, polyethylene glycol, protoplast-fusion, etc.

b) DNA fragments corresponding to genes can be cloned into expression vectors for *E. coli*, yeast or mammalian cells and the resultant proteins purified.

c) The provival DNA can be "shot-gunned" (fragmented) into procaryotic expression vectors to generate fusion polypeptides.

Recombinants, producing antigenically competent fusion proteins,. can be identified by simply screening the recombinants with antibodies against LAV_{MAL} antigens. Particular reference in this respect is made to those portions of the genome of LAV_{MAL} which, in the figures, are shown to belong to open reading frames and which encode the products having the polypeptidic backbones shown.

Different polypeptides which appear in figures 7A to 7I are still further provided. Methods disclosed in European application 0 178 978 and in PCT application PCT/EP 85/00548, filed Oct.18, 1985, are applicable for the production of such peptides from LAV_{MAL}. In this regard, polypeptides are provided containing sequences in common with polypeptides comprising antigenic determinants included in the proteins encoded and expressed by the LAV_{MAL} genome. Means are also provided for the detection of proteins of LAV_{MAL}, particularly for the diagnosis of AIDS or pre-AIDS or, to the contrary, for the detection of antibodies against LAV_{MAL} or its proteins, particularly in patients afflicted with AIDS or pre-AIDS or more generally in asymtomatic carriers and in blood-related products. Further provided are immunogenic polypeptides and more particularly protective polypeptides for use in the preparation of vaccine compositions against AIDS or related syndroms.

Yet further provided are polypeptide fragments having lower molecular weights and having peptide sequences or fragments in common with those shown in figures 7A to 7I. Fragments of smaller sizes can be obtained by resorting to known techniques, for instance, by cleaving the original larger polypeptide by enzymes

capable of cleaving it at specific sites. By way of examples may be mentioned the enzyme of Staphylococcyus aureus V8, α-chymotrypsine, "mouse sub-maxillary gland protease" marketed by the Boehringer company, Vibrio alginolyticus chemovar iophagus collagenase, which specifically recognizes the peptides Gly-Pro, Gly-Ala, etc.

Other features of this invention will appear in the following disclosure of data obtained starting from LAV_{MAL}, in relation to the drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

- Figs. 1A and 1B provide comparative restriction maps of the genomas of LAV_{MAL} as compared to LAV_{ELI} (Applicants' related new LAV virus which is the subject of their copending application, filed herewith) and LAV_{BRU} (a known LAV isolate deposited at the Collection Nationale des Cultures de Micro-organismes (hereinafter "CNCM") of the Pasteur Institute, Paris, France under No. I-232 on July 15, 1983) ;
- Fig. 2 shows comparative maps setting forth the relative positions of the open reading frames of the above genomas ;
- Figs. 3A-3F (also designated generally hereinafter "fig. 3") indicate the relative correspondance between the proteins (or glycoproteins) encoded by the open reading frames, whereby amino acid residues of protein sequences of LAV_{MAL} are in vertical alignment with corresponding amino acid residues (numbered) of corresponding or homologous proteins or glycoproteins of LAV_{BRU}, as well as LAV_{ELI} and ARV 2.
- Figs. 4A-4B (also designated generally hereinafter "fig. 4") provide tables quantitating the sequence divergence between homologous proteins of LAV_{BRU}, LAV_{ELI}

and LAV_{MAL} ;

- Fig. 5 shows diagrammatically the degree of divergence of the different virus envelope proteins ;
- Figs. 6A and 6B ("Fig. 6" when consulted together) render apparent the direct repeats which appear in the proteins of the different AIDS virus isolates.
- Figs. 7A-7I show the full nucleotidic sequences of LAV_{MAL}.

DETAILED DESCRIPTION OF THE INVENTION

CHARACTERIZATION AND MOLECULAR CLONING OF AN AFRICAN ISOLATE.

The different AIDS virus isolates concerned are designated by three letters of the patients name, LAV_{BRU} referring to the prototype AIDS virus isolated in 1983 from a French homosexual patient with LAS and thought to have been infected in the USA in the preceding years [Barré-Sinoussi et al., 1983]. LAV_{MAL} was recovered in 1985 from a 7-year old boy from Zaire. Related LAV_{ELI} was recovered in 1983 from a 24-year old woman with AIDS from Zaire. Recovery and purification of the LAV_{MAL} virus were performed according to the method disclosed in European Patent Application 84 401834/138 667 filed on September 9, 1984.

LAV_{MAL} is indistinguishable from the previously characterized isolates by its structural and biological properties in vitro. Virus metabolic labelling and immune precipitation by patient MAL sera, as well as reference sera, showed that the proteins of LAV_{MAL} had the same molecular weight (hereinafter "MW") as, and cross-reacted immunologically with those of, prototype AIDS virus (data not shown) of the LAV-1 class.

Reference is again made to European Application 178 978 and International Application PCT/EP

85/00548 as concerns the purification, mapping and sequencing procedures used herein. See also the discussion under the headings "Experimental Procedures" and "Significance of the Figures" hereinafter.

Primary restriction enzyme analysis of LAV_{MAL} genome was done by southern blot with total DNA derived from acutely infected lymphocytes, using cloned LAV_{BRU} complete genome as probe. Overall cross-hybridization was observed under stringent conditions, but the restriction profile of the Zairian isolate was clearly different. Phage lambda clones carrying the complete viral genetic information were obtained and further characterized by restriction mapping and nucleotide sequence analysis. A clone (hereinafter "M-H11") was obtained by complete HindIII restriction of DNA from LAV_{MAL}-infected cells, taking advantage of the existence of a unique HindIII site in the long terminal repeat (hereinafter "LTR"). M-H11 is thus probably derived from unintegrated viral DNA since that species was at least ten times more abundant than integrated provirus.

Figure 1B gives a comparaison of the restriction maps derived from the nucleotide sequences of LAV_{ELI}, LAV_{MAL} and prototype LAV_{BRU}, as well as from three other Zairian isolates (hereinafter "Z1", "Z2", and "Z3" respectively) previously mapped for seven restriction enzymes [Benn et al., 1985]. Despite this limited number, all of the profiles are clearly different (out of the 23 sites making up the map of LAV_{BRU}, only seven are present in all six maps presented), confirming the genetic polymorphism of the AIDS virus. No obvious relationship is apparent between the five Zairian maps, and all of their common sites are also found in LAV_{BRU}.

Conservation of the genetic organization.

The genetic organization of LAV_{MAL} as deduced

from the complete nucleotide sequences of its cloned genome is identical to that found in other isolates, i.e., 5'gag-pol-central region-env-F3'. Most noticeable is the conservation of the "central region" (fig. 2), located between the pol and env genes, which is composed of a series of overlapping open reading frames (hereinafter "orf") previously designated Q, R, S, T, and U in the ovine lentivirus visna [Sonigo et al., 1985]. The product of orf S (also designated "tat") is implicated in the transactivation of virus expression [Sodroski et al., 1985 ; Arya et al., 1985] ; the biological role of the product of orf Q (also designated "sor" or "orf A") is still unknown [Lee et al., 1986 ; Kang et al., 1986]. Of the three other orfs, R, T, and U, only orf R is likely to be a seventh viral gene, for the following reasons : the exact conservation of its relative position with respect to Q and S (fig. 2), the ponstant presence of a possible splice acceptor and of a consensus AUG initiator codon, its similar codon usage with respect to viral genes, and finally the fact that the variation of its protein sequence within the different isolates is comparable to that of gag, pol and Q (see fig. 4).

Also conserved are the sizes of the U3, R and U5 elements of the LTR (data not shown), the location and sequence of their regulatory elements such as TATA box and AATAAA polyadenylation signal, and their flanking sequences, i.e., primer binding site (hereinafter "PBS") complementary to 3' end of $tRNA^{LYS}$ and polypurine tract (hereinafter "PPT"). Most of the genetic variability within the LTR is located in the 5' half of U3 (which encodes a part of orf F) while the 3' end of U3 and R, which carry most of the cis-acting regulatory elements, promoter, enhancer and trans-activating factor receptor [Rosen et al., 1985],

as well as the U5 element, are well-conserved.

Overall, it clearly appears that this Zairian isolate, LAV_{MAL}, is the same type of retrovirus as the previously sequenced isolates of American or European origin.

Variability of the viral proteins.

Despite their identical genetic organization, the LAV_{ELI} and LAV_{MAL} shows substantial differences in the primary structure of their proteins. The amino acid sequences of LAV_{ELI} and LAV_{MAL} proteins are presented in figures 3A-3F, aligned with those of LAV_{BRU} and ARV 2. Their divergence was quantified as the percentage of amino acids substitutions in two-by-two alignments (Fig. 4). The number of insertions and deletions that had to be introduced in each of these alignments has also been scored.

Three general observations can be made. First, the protein sequences of the LAV_{ELI} and LAV_{MAL} are more divergent from LAV_{BRU} than are those of HTLV-3 and ARV 2 (Fig. 4A) ; similar results are obtained if ARV 2 is taken as reference (not shown). The range of genetic polymorphism between isolates of the AIDS virus is considerably greater than previously observed. Second, our two sequences confirm that the envelope is more variable than the gag and pol genes. Here again, the relatively small difference observed between the env of LAV_{BRU} and HTLV-3 appears as an exception. Third, the mutual divergence of the LAV_{ELI} and LAV_{MAL} (Fig. 4B) is comparable to that between LAV_{BRU} and either of them; as far as we can extrapolate from only three sequenced isolates from the USA and Europe and two (LAV_{ELI} and LAV_{MAL}) from Africa, this is indicative of a wider evolution of the AIDS virus in Africa.

<u>gag and pol</u> : Their greater degree of conservation compared to the envelope is consistent with their

encoding important structural or enzymatic activities. Of the three mature gag proteins, the p25 which was the first recognized immunogenic protein of LAV [Barré-Sinoussi et al., 1983] is also the better conserved (fig. 3). In gag and pol, differences between isolates are principally due to point mutations, and only a small number of insertional or deletional events is observed. Among these, we must note the presence in the overlapping part of gag and pol of LAV_{BRU} of an insertion of 12 amino acids (AA) which is encoded by the second copy of a 36 bp direct repeat present only in this isolate and in HTLV-3. This duplication was omitted because of a computing error in the published sequence of LAV_{BRU} (position 1712, Wain-Hobson et al., 1985) but was indeed present in the HTLV-3 sequences [Ratner et al., 1985 ; Muesing et al., 1985].

<u>env</u> : Three segments can be distinguished in the envelope glycoprotein precursor [Allan et al., 1985 ; Montagnier et al., 1985 ; DiMarzoVeronese et al., 1985]. The first is the signal peptide (positions 1-33 in fig. 3), and its sequence appears as variable ; the second segment (pos. 34-530) forms the outer membrane protein (hereinafter "OMP" or "gp110") and carries most of the genetic variations, and in particular almost all of the numerous reciprocal insertions and deletions ; the third segment (531-877) is separated from the OMP by a potential cleavage site following a constant basic stretch (Arg-Glu-Lys-Arg) and forms the transmembrane protein (hereinafter "TMP" or "gp 41") responsible for the anchorage of the envelope glycoprotein in the cellular membrane. A better conservation of the TMP than the OMP has also been observed between the different murine leukemia viruses (hereinafter "MLV") [Koch et al., 1983] and could be due to structural constraints.

From the alignment of figure 3 and the

graphical representation of the envelope variability shown in figure 5, we clearly see the existence of conserved domains, with little or no genetic variation, and hypervariable domains, in which even the alignment of the different sequences is very difficult, because of the existence of a large number of mutations and of reciprocal insertions and deletions. We have not included the sequence of the envelope of the HTLV-3 isolate since it so close to that of LAV_{BRU} (cf. fig. 4), even in the hypervariable domains, that it did not add anything to the analysis. While this graphical representation will be refined by more sequence data, the general profile is already apparent, with three hypervariable domains (Hyl, 2 and 3) all being located in the OMP and separated by three well-conserved stretches (residues 37-130, 211-289, and 488-530 of fig. 3 alignment) probably associated with important biological functions.

In spite of the extreme genetic variability, the folding pattern of the envelope glycoprotein is probably constant. Indeed the position of virtually all of the cysteine residues is conserved within the different isolates (fig. 3 and 5), and the only three variable cysteines fall either in the signal peptide or in the very C-terminal part of the TMP. The hypervariable domains of the OMP are bounded by conserved cysteines, suggesting that they may represent loops attached to the common folding pattern. Also the calculated hydropathic profiles [Kyte and Doolittle, 1982] of the different envelope proteins are remarkably conserved (not shown).

About half of the potential N-glycosylation sites, Asn-X-Ser/Thr, found in the envelopes of the Zairian isolates map to the same positions in LAV_{BRU} (17/26 for LAV_{ELI} and 17/28 for LAV_{MAL}). The other sites appear to fall within variable domains of env,

suggesting the existence of differences in the extent of envelope glycosylation between different isolates.

Other viral proteins : Of the three other identified viral proteins, the p27 encoded by orf F, 3' of env [Allan et al., 1985b] is the most variable (fig. 4). The proteins encoded by orfs Q and S of the central region are remarkable by their absence of insertions/deletions. Surprisingly, a high frequency of amino acids substitutions, comparable to that observed in env, is found for the product of orf S (trans-activating factor). On the other hand, the protein encoded by orf Q is no more variable than gag. Also noticeable is the lower variation of the proteins encoded by the central regions of LAV_{ELI} and LAV_{MAL}.

With the availability of the complete nucleotide sequence from five independant isolates, some general features of the AIDS virus' genetic variability are now emerging. Firstly, its principal cause is point mutations which very often result in amino acid substitutions and which are more frequent in the 3' part of the genome (orf S, env and orf F). Like all RNA viruses, the retroviruses are thought to be highly subject to mutations caused by errors of the RNA polymerases during their replication, since there is no proofreading, of this step [Holland et al., 1982 ; Steinhauer and Holland, 1986].

Another source of genetic diversity is insertions/deletions. From the figure 3 alignments, insertional events seem to be implicated in most of the cases, since otherwise deletions should have occurred in independant isolates at precisely the same locations. Furthermore, upon analyzing these insertions, we have observed that they most often represent one of the two copies of a direct repeat (fig. 6). Some are perfectly conserved like the 36 bp repeat in the gag-pol overlap

of LAV_{BRU} (fig. 6-a) ; others carry point mutations resulting in amino acid substitutions, and as a consequence, they are more difficult to observe, though clearly present, in the hypervariable domains of env (cf. fig. 6-g and -h). As noted for point mutations, env gene and orf F also appear as more susceptible to that form of genetic variation than the rest of the genome. The degree of conservation of these repeats must be related to their date of occurrence in the analyzed sequences : the more degenerated, the more ancient. A very recent divergence of LAV_{BRU} and HTLV-3 is suggested by the extremely low number of mismatched AA between their homologous proteins. However, one of the LAV_{BRU} repeats (located in the Hyl domain of env, fig. 6-f) is not present in HTLV-3, indicating that this generation of tandem repeats is a rapid source of genetic diversity. We have found no traces of such a phenomenon, even when comparing very closely related viruses, such as the Mason-Pfizer monkey virus (hereinafter "MPMV") [Sonigo et al., 1986], and an immunosuppressive simian virus (hereinafter "SRV-1") [Power et al., 1986]. Insertion or deletion of one copy of a direct repeat have been occasionally reported in mutant retroviruses [Shimotohno and Temin, 1981 ; Darlix, 1986], but the extent to which we observe this phenomenon is unprecedented. The molecular basis of these duplications is unclear, but could be the "copy-choice" phenomenon, resulting from the diploidy of the retroviral genome [Varmus and Swanstrom, 1984 ; Clark and Mak, 1983]. During the synthesis of the first-strand of the viral DNA, jumps are known to occur from one RNA molecule to another, especially when a break or a stable secondary structure is present on the template; an inaccurate re-initiation on the other RNA template could result in the generation (or the elimination) of a short direct repeat.

Genetic variability and subsequent antigenic modifications have often been developed by micro-organisms as a means for avoiding the host's immune response, either by modifying their epitopes during the course of the infection, as in trypanosomes [Borst and Cross, 1982], or by generating a large repertoire of antigens, as observed in influenza virus [Webster et al., 1982]. As the human AIDS virus is related to animal lentiviruses [Sonigo et al., 1985 ; Chiu et al., 1985], its genetic variability could be a source of antigenic variation, as can be observed during the course of the infection by the ovine lentivirus visna [Scott et al., 1979 ; Clements et al., 1980] or by the equine infectious anemia virus (hereinafter "EIAV") [Montelaro et al., 1984]. However, a major discrepancy with these animal models is the extremely low, and possibly nonexistant, neutralizing activity of the sera of individuals infected by the AIDS virus, whether they are healthy carriers, displaying minor symptoms, or afflicted with AIDS [Weiss et al., 1985 ; Clavel et al., 1985]. Furthermore, even for the visna virus the exact role of antigenic variation in the pathogenesis is unclear [Thormar et al., 1983 ; Lutley et al., 1983]. We rather believe that genetic variation represents a general selective advantage for lentiviruses by allowing an adaptation to different environments, for example by modifying their tissue or host tropisms. In the particular case of the AIDS virus, rapid genetic variations are tolerated, especially in the envelope. This could allow the virus to become adapted to different "micro-environments" of the membrane of their principal target cells, namely the T4 lymphocytes. These "micro-environments" could result from the immediate vicinity of the virus receptor to polymorphic surface proteins, differing either between individuals or between clones of

lymphocytes.

Conserved domains in the AIDS virus envelope

Since the proteins of most of the isolates are antigenically cross-reactive, the genotypic differences do not seem to affect the sensitivity of actual diagnostic tests, based upon the detection of antibodies to the AIDS virus and using purified virions as antigens. They nevertheless have to be considered for the development of the "second-generation" tests, that are expected to be more specific, and will use smaller synthetic or genetically-engineered viral antigens. The identification of conserved domains in the highly immunogenic envelope glycoprotein and the core structural proteins (gag) is very important for these tests. The conserved stretch found at the end of the OMP and the beginning of the TMP (490-620, fig. 3) could be a good candidate, since a bacterial fusion protein containing this domain was well-detected by AIDS patients' sera [Chang et al., 1985].

The envelope, specifically the OMP, mediates the interaction between a retrovirus and its specific cellular receptor [DeLarco and Todaro, 1976 ; Robinson et al., 1980]. In the case of the AIDS virus, _in vitro_ binding assays have shown the interaction of the envelope glycoprotein gp110 with the T4 cellular surface antigen [McDougal et al., 1986], already thought to be closely associated with the virus receptor [Klatzmann et al., 1984 ; Dagleish et al., 1984]. Identification of the AIDS virus envelope domains that are responsible for this interaction (receptor-binding domains) appears to be fundamental for understanding of the host-viral interactions and for designing a protective vaccine, since an immune response against these epitopes could possibly elicit neutralizing antibodies. As the AIDS

virus receptor is at least partly formed of a constant structure, the T4 antigen, the binding site of the envelope is unlikely to be exclusively encoded by domains undergoing drastic genetic changes between isolates, even if these could be implicated in some kind of an "adaptation". One or several of the conserved domains of the OMP (residues 37-130, 211-289, and 488-530 of fig. 3 alignment), brought together by the folding of the protein, must play a part in the virus-receptor interaction, and this can be explored with synthetic or genetically-engineered peptides derived from these domains, either by direct binding assays or indirectly by assaying the neutralizing activity of specific antibodies raised against them.

<u>African AIDS viruses</u>

Zaire and the neighbouring countries of Central Africa are considered as an area endemic with the AIDS virus infection, and the possibility that the virus has emerged in Africa has became a subject of intense controversy (see Norman, 1985). From the present study, it is clear that the genetic organization of Zairian isolates is the same as that of american isolates, thereby indicating a common origin. The very important sequence differences observed between the proteins are consistent with a divergent evolutionary process. In addition, the two African isolates are mutually more divergent than the American isolates already analyzed ; as far as that observation can be extrapolated, it suggests a longer evolution of the virus in Africa and is also consistent with the fact that a larger fraction of the population is exposed than in developed countries.

A novel human retrovirus with morphology and biologocal properties (cytopathogenicity, T4 tropism)

similar to those of LAV, but nevertheless clearly genetically and antigenically distinct from it, was recently isolated from two patients with AIDS originating from Guinea Bissau, West-Africa [Clavel et al., 1986]. In neighboring Senegal, the population was seemingly exposed to a retrovirus also distinct from LAV but apparently non-pathogenic [Barin et al., 1985 ; Kanki et al., 1986]. Both of these novel African retroviruses seem to be antigenically related to the simian T-cell lymphotropic virus (hereinafter "STLV-III") shown to be widely present in healthy African green monkeys and other simian species [Kanki et al. 1985]. This raises the possibility of a large group of African primate lentiviruses, ranging from the apparently non-pathogenic simian viruses to the LAV-type viruses. Their precise relationship will only be known after their complete genetic characterization, but it is already very likely that they have evolved from a common progenitor. The important genetic variability we have observed between isolates of the AIDS virus in Central Africa is probably a hallmark of this entire group and may account for the apparently important genetic divergence between its members (loss of cross-antigenicity in the envelopes). In this sense, the conservation of the tropism for the T4 lymphocytes suggests that it is a major advantage aquired by these retroviruses.

EXPERIMENTAL PROCEDURES

Virus isolation

LAV_{MAL} was isolated from the peripheral blood lymphocytes of the patient as described [Barré-Sinoussi et al., 1983]. Briefly, the lymphocytes were fractionated and co-cultivated with phytohaemagglutinin-stimulated normal human lymphocytes in the presence of

interleukin 2 and anti-alpha interferon serum. Viral production was assessed by cell-free reverse transcriptase (hereinafter "RT") activity assay in the cultures and by electron microscopy.

Molecular cloning

Normal donor lymphocytes were acutely infected (10^4 cpm of RT activity/10^6 cells) as described [Barré-Sinoussi et al., 1983], and total DNA was extracted at the beginning of the RT activity peak. A lambda library using the L47-1 vector [Loenen and Brammar, 1982] was constructed by partial HindIII digestion of the DNA as already described [Alizon et al., 1984]. DNA from infected cells was digested to completion with HindIII, and the 9-10kb fraction was selected on 0.8 % low melting point agarose gel and ligated into L47-1 HindIII arms. About 2.10^5 plaques for LAV_{MAL}, obtained by in vitro packaging (Amersham), were plated on E. coli LA101 and screened in situ under stringent conditions,using the 9 kb SacI insert of the clone lambda J19 [Alizon et al., 1984] carrying most of the LAV_{BRU} genome as probe. Clones displaying positive signals were plaque-purified and propagated on E. coli C600 recBC, and the recombinant phage M-H11 carrying the complete genetic information of LAV_{MAL} was further characterized by restriction mapping.

Nucleotide sequence strategy

Viral fragments derived from M-H11 were sequenced by the dideoxy chain terminator procedure [Sanger et al., 1977] after "shotgun" cloning in the M13mp8 vector [Messing and Viera, 1982] as previously described [Sonigo et al., 1985]. The viral genome of LAV_{MAL} is 9229 nucleotides long as shown in figs. 7A-7I. Each nucleotide of LAV_{MAL} was determined from more than 5 independent clones on average.

SIGNIFICANCE OF THE FIGURES

Figure 1 contains an analysis of AIDS virus isolates, showing:

A/ Restriction maps of the inserts of phage lambda clones derived from cells infected with LAV_{ELI} (hereinafter "E-H12") and with LAV_{MAL} (M-H11). The schematic genetic organization of the AIDS virus has been drawn above the maps. The LTRs are indicated by solid boxes. Restriction sites are indicated as follows: A:Aval; B:BamHI; Bg:BgIII; E:EcoRI; H:HindIII; Hc:HincII; K:KpnI; N:NdeI; P:PstI; S:SacI; and X:XbaI. Asterisks indicate the HindIII cloning sites in lambda L47-1 vector.

B/ A comparison of the sites for seven restriction enzymes in six isolates : the prototype AIDS virus LAV_{BRU}, LAV_{MAL} and LAV_{ELI} ; and Z1, Z2 and Z3. Restriction sites are represented by the following symbols vertically aligned wih the symbols in fig. 1A: ●: BgIII; * :EcoRI; ▽ :HincII; ▼:HindIII; ◆:KpnI; ◇ :NdeI; and *o* :SacI.

Figure 2 shows the genetic organization of the central region in AIDS virus isolates. Stop codons in each phase are represented as vertical bars. Vertical arrows indicate possible AUG initiation codons. Splice acceptor (A) and donor (D) sites identified in subgenomic viral mRNA [Muesing et al., 1985] are shown below the graphic of LAV_{BRU}, and corresponding sites in LAV_{ELI} and LAV_{MAL} are indicated. PPT indicates the repeat of the polypurine tract flanking the 3'LTR. As observed in LAV_{BRU} [Wain-Hobson et al., 1985], the PPT is repeated 256 nucleotides 5' to the end of the pol gene in both the LAV_{ELI} and LAV_{MAL} sequences, but this repeat is degenerated at two positions in LAV_{ELI}.

Figure 3 shows an alignment of the protein sequences of four AIDS virus isolates. Isolate LAV_{BRU} [Wain-Hobson et

al., 1985] is taken as reference ; only differences with LAV_{BRU} are noted for ARV 2 [Sanchez-Pescador et al., 1985] and the two Zairian isolates LAV_{MAL} and LAV_{ELI}. A minimal number of gaps (-) were introduced in the alignments. The NH_2-termini of $p25^{gag}$ and $p18^{gag}$ are indicated [Sanchez-Pescador, 1985]. The potential cleavage sites in the envelope precursor [Allan et al., 1985a ; diMarzoVeronese, 1985] separating the signal peptide (hereinafter "SP"), OMP and TMP are indicated as vertical arrows ; conserved cysteines are indicated by black circles and variable cysteines are boxed. The one letter code for each amino acid is as follows: A:Ala ; C:Cys ; D:Asp ; E:Glu ; F:Phe ; G:Gly ; H :His ; I:Ile ; K:Lys ; L:Leu ; M:Met ; N:Asn ; P:Pro ; Q:Gln ; R:Arg ; S:Ser ; T:Thr ; V:Val; W:Trp ; Y:Tyr.

Figure 4 shows a quantitation of the sequence divergence between homologous proteins of different isolates. Part A of each table gives results deduced from two-by-two alignments using the proteins of LAV_{BRU} as reference, part B, those of LAV_{ELI} as reference. Sources: Muesing et al., 1985 for HTLV-3 ; Sanchez-Pescador et al., 1985 for ARV 2 and Wain-Hobson et al., 1985 for LAV_{BRU}. For each case in the tables, the size in amino acids of the protein (calculated from the first methionine residue or from the beginning of the orf for pol) is given at the upper left part. Below are given the number of deletions (left) and insertions (right) necessary for the alignment. The large numbers in bold face represent the percentage of amino acids substitutions (insertions/deletions being excluded). Two by two alignments were done with computer assistance [Wilburg and Lipman, 1983], using a gag penalty of 1, K-tuple of 1, and window of 20, except for the hypervariable domains of env, where the number of gaps was made minimum, and which are essentially aligned as shown in fig. 3. The sequence of

the predicted protein encoded by orf R of HTLV-3 has not been compared because of a premature termination relative to all other isolates.

Figure 5 shows the variability of the AIDS virus envelope protein. For each position x of the alignment of <u>env</u> (Fig. 3), variability V(x) was calculated as:

V(x) = number of different amino-acids at position x/ frequency of the most abundant amino acid at position x. Gaps in the alignments are considered as another amino acid. For an alignment of 4 proteins, V(x) ranges from 1 (identical AA in the 4 sequences) to 16 (4 different AA). This type of representation has previously been used in a compilation of the AA sequence of immunoglobulins variable regions [Wu and Kabat, 1970]. Vertical arrows indicate the cleavage sites ; asterisks represent potential N-glysosylation sites (N-X-S/T) conserved in all three four solates ; black triangles represent conserved cysteine residues. Black lozanges mark the three major hydrophobic domains: OMP, TMP and SP; and the hypervariable domains: Hyl, 2 and 3.

Figure 6 shows the direct repeats in the proteins of different AIDS virus isolates. These examples are derived from the aligned sequences of gag (a, b), F (c,d) and env (e, f, g, h) shown in figure 3. The two elements of the direct repeat are boxed, while degenerated positions are underlined.

Figures 7A-7I show the complete cDNA sequence of LAV_{MAL} of this invention.

The invention thus pertains more specifically to the proteins, glycoproteins and other polypeptides including the polypeptidic structures shown in the figures 1-7. The first and last amino acid residues of these proteins, glycoproteins and polypeptides carry numbers computed from a first amino acid of the open-reading frames concerned, although these numbers do

not correspond exactly to those of the LAV_{MAL} proteins concerned, rather to the corresponding proteins of the LAV_{BRU} or sequences shown in figs. 3A, 3B and 3C. Thus a number corresponding to a "first amino acid residue" of a LAV_{MAL} protein corresponds to the number of the first amino-acyl residue of the corresponding LAV_{BRU} protein which, in any of figs. 3A, 3B or 3C, is in direct alignment with the corresponding first amino acid of the LAV_{MAL} protein. Thus the sequences concerned can be read from figs. 7A-7I to the extent where they do not appear with sufficient clarity from Figs. 3A-3F.

The preferred protein sequences of this invention extend between the corresponding "first" and "last" amino acid residues. Also preferred are the protein(s)- or glycoprotein(s)-portions including part of the sequences which follow :

OMP or gp110 proteins, including precursors :

1 to 530

OMP or gp110 without precursor :

34-530

Sequence carrying the TMP or gp41 protein :

531-877, particularly

680-700

well conserved stretches of OMP :

37-130,

211-289 and

488-530

well conserved stretch found at the end of the OMP and the beginning of TMP :

490-620.

Proteins containing or consisting of the "well conserved stretches" are of particular interest for the production of immunogenic compositions and (preferably in relation to the stretches of the env protein) of vaccine compositions against the LAV-1 viruses.

The invention concerns more particularly all the DNA fragments which have been more specifically referred to in the drawings and which correspond to open reading frames. It will be understood that one skilled in the art will be able to obtain them all, for instance by cleaving an entire DNA corresponding to the complete genome of LAV_{MAL}, such as by cleavage by a partial or complete digestion thereof with a suitable restriction enzyme and by the subsequent recovery of the relevant fragments. The DNA disclosed above can be resorted to also as a source of suitable fragments. The techniques disclosed in PCT application for the isolation of the fragments which can then be included in suitable plasmids are applicable here too. Of course, other methods can be used, some of which have been examplified in European Application No. 178,978, filed September 17, 1985. Reference is for instance made to the following methods:

a) DNA can be transfected into mammalian cells with appropriate selection markers by a variety of techniques, such as calcium phosphate precipitation, polyethylene glycol, protoplast-fusion, etc.

b) DNA fragments corresponding to genes can be cloned into expression vectors for E. coli, yeast- or mammalian cells and the resultant proteins purified.

c) The provival DNA can be "shot-gunned" (fragmented) into procaryotic expression vectors to generate fusion polypeptides. Recombinants, producing antigenically competent fusion proteins, can be identified by simply screening the recombinants with antibodies against LAV antigens.

The invention further refers to DNA recombinants, particularly modified vectors, including any of the preceding DNA sequences adapted to transform corresponding microorganisms or cells, particularly eucaryotic

cells such as yeasts, for instance <u>Saccharomyces cerevisiae</u>, or higher eucaryotic cells, particularly cells of mammals, and to permit expression of said DNA sequences in the corresponding microorganisms or cells. General methods of that type have been recalled in the abovesaid PCT international patent aplication PCT/EP 85/00548, filed October 18, 1985.

More particularly the invention relates to such modified DNA recombinant vectors modified by the abovesaid DNA sequences and which are capable of transforming higher eucaryotic cells particularly mammalian cells. Preferably, any of the abovesaid sequences are placed under the direct control of a promoter contained in said vectors and recognized by the polymerases of said cells, such that the first nucleotide codons expressed correspond to the first triplets of the above-defined DNA sequences. Accordingly, this invention also relates to the corresponding DNA fragments which can be obtained from the genome of LAV_{MAL} or its cDNA by any appropriate method. For instance, such a method comprises cleaving said LAV_{MAL} genome or its cDNA by restriction enzymes preferably at the level of restriction sites surrounding said fragments and close to the opposite extremities respectively thereof, recovering and identifying the fragments sought according to sizes, if need be checking their restriction maps or nucleotide sequences (or by reaction with monoclonal antibodies specifically directed against epitopes carried by the polypeptides encoded by said DNA fragments), and further if need be, trimming the extremities of the fragment, for instance by an exonucleolytic enzyme such as Bal31, for the purpose of controlling the desired nucleotid-sequences of the extremities of said DNA fragments or, conversely, repairing said extremities with Klenow enzyme and possibly ligating the latter to synthetic

polynucleotide fragments designed to permit the reconstitution of the nucleotide extremities of said fragments. Those fragments may then be inserted in any of said vectors for causing the expression of the corresponding polypeptide by the cell transformed therewith. The corresponding polypeptide can then be recovered from the transformed cells, if need be after lysis thereof, and purified by methods such as electrophoresis. Needless to say, all conventional methods for performing these operations can be resorted to.

The invention also relates more specifically to cloned probes which can be made starting from any DNA fragment according to this invention, thus to recombinant DNAs containing such fragments, particularly any plasmids amplifiable in procaryotic or eucaryotic cells and carrying said fragments. Using the cloned DNA fragments as a molecular hybridization probe - either by labelling with radionucleotides or with fluorescent reagents - LAV virion RNA may be detected directly in the blood, body fluids and blood products (e.g. of the antihemophylic factors such as Factor VIII concentrates) and vaccines (e.g., hepatitis B vaccine). It has already been shown that whole virus can be detected in culture supernatants of LAV producing cells. A suitable method for achieving that detection comprises immobilizing virus on a support (e.g., a nitrocellulose filter), disrupting the virion and hybridizing with labelled (radiolabelled or "cold" fluorescent- or enzyme-labelled) probes. Such an approach has already been developed for Hepatitis B virus in peripheral blood [SCOTTO J. et al. Hepatology (1983), 3, 379-384].

Probes according to the invention can also be used for rapid screening of genomic DNA derived from the tissue of patients with LAV related symptoms, to see if the proviral DNA or RNA present in host tissue and other

tissues can be related to that of LAV_{MAL}.

A method which can be used for such screening comprises the following steps : extraction of DNA from tissue, restriction enzyme cleavage of said DNA, electrophoresis of the fragments and Southern blotting of genomic DNA from tissues, subsequent hybridization with labelled cloned LAV proviral DNA. Hybridization in situ can also be used.

Lymphatic fluids and tissues and other non-lymphatic tissues of humans, primates and other mammalian species can also be screened to see if other evolutionnary related retrovirus exist. The methods referred to hereinabove can be used, although hybridization and washings would be done under non-stringent conditions.

The DNAs or DNA fragments according to the invention can be used also for achieving the expression of viral antigens of LAV_{MAL} for diagnostic purposes.

The invention relates generally to the polypeptides themselves, whether synthesized chemically, isolated from viral preparations or expressed by the different DNAs of the invention, particularly by the ORFs or fragments thereof in appropriate hosts, particularly procaryotic or eucaryotic hosts, after transformation thereof with a suitable vector previously modified by the corresponding DNAs.

More generally, the invention also relates to any of the polypeptide fragments (or molecules, particularly glycoproteins having the same polypeptidic backbone as the polypeptides mentioned hereinabove) bearing an epitope characteristic of a protein or glycoprotein of LAV_{MAL}, which polypeptide or molecule then has N-terminal and C-terminal extremities respectively either free or, independently from each other, covalently bonded to amino acids other than those which are

normally associated with them in the larger polypeptides or glycoproteins of the LAV virus, which last mentioned amino acids are then free or belong to another polypeptidic sequence. Particularly, the invention relates to hybrid polypeptides containing any of the epitope-bearing-polypeptides which have been defined more specifically hereinabove, recombined with other polypeptides fragments normally foreign to the LAV proteins, having sizes sufficient to provide for an increased immunogenicity of the epitope-bearing-polypeptide yet, said foreign polypeptide fragments either being immunogenically inert or not interfering with the immunogenic properties of the epitope-bearing-polypeptide.

Such hybrid polypeptides, which may contain from 5 up to 150, even 250 amino acids, usually consist of the expression products of a vector which contained <u>ab initio</u> a nucleic acid sequence expressible under the control of a suitable promoter or replicon in a suitable host, which nucleic acid sequence had however beforehand been modified by insertion therein of a DNA sequence encoding said epitope-bearing-polypeptide.

Said epitope-bearing-polypeptides, particularly those whose N-terminal and C-terminal amino acids are free, are also accessible by chemical synthesis according to technics well known in the chemistry of proteins.

The synthesis of peptides in homogeneous solution and in solid phase is well known. In this respect, recourse may be had to the method of synthesis in homogeneous solution described by Houbenweyl in the work entitled "Methoden der Organischen Chemie" (Methods of Organic Chemistry) edited by E. WUNSCH., vol. 15-I and II, THIEME, Stuttgart 1974. This method of synthesis consists of successively condensing either the successive amino acids in twos, in the appropriate order or successive peptide fragments previously available or

formed and containing already several amino-acyl residues in the appropriate order respectively. Except for the carboxyl and aminogroups which will be engaged in the formation of the peptide bonds, care must be taken to protect beforehand all other reactive groups borne by these amino-acyl groups or fragments. However, prior to the formation of the peptide bonds, the carboxyl groups are advantageously activated, according to methods well known in the synthesis of peptides. Alternatively, recourse may be had to coupling reactions bringing into play conventional coupling reagents, for instance of the carbodiimide type, such as 1-ethyl-3-(3-dimethyl-amino-propyl)-carbodiimide. When the amino acid group used carries an additional amine group (e.g., lysine) or another acid function (e.g., glutamic acid), these groups may be protected by carbobenzoxy or t-butyloxycarbonyl groups, as regards the amine groups, or by t-butylester groups, as regards the carboxylic groups. Similar procedures are available for the protection of other reactive groups. for example, an -SH group (e.g., in cysteine) can be protected by an acetamidomethyl or paramethoxybenzyl group.

In the case of a progressive synthesis, amino acid by amino acid, the synthesis starts preferably with the condensation of the C-terminal amino acid with the amino acid which corresponds to the neighboring amino-acyl group in the desired sequence and so on, step by step, up to the N-terminal amino acid. Another preferred technique which can be used is that described by R.D. Merrifield in "Solid Phase Peptide Synthesis" [J. Am. Chem. Soc., 45, 2149-2154]. In accordance with the Merrifield process, the first C-terminal amino acid of the chain is fixed to a suitable porous polymeric resin, by means of its carboxylic group, the amino group of the

amino acid then being protected, for example by a t-butyloxycarbonyl group. When the first C-terminal amino acid is thus fixed to the resin, the protective group of the amine group is removed by washing the resin with an acid, i.e., trifluoroacetic acid, when the protective group of the amine group is a t-butyloxycarbonyl group. Then, the carboxylic group of the second amino acid, which is to provide the second amino-acyl group of the desired peptidic sequence, is coupled to the deprotected amine group of the C-terminal amino acid fixed to the resin. Preferably, the carboxyl group of this second amino acid has been activated, for example by dicyclohexyl-carbodiimide, while its amine group has been protected, for example by a t-butyloxycarbonyl group. The first part of the desired peptide chain, which comprises the first two amino acids, is thus obtained. As previously, the amine group is then deprotected, and one can further proceed with the fixing of the next amino-acyl group and so forth until the whole peptide sought is obtained. The protective groups of the different side groups, if any, of the peptide chain so formed can then be removed. The peptide sought can then be detached from the resin, for example by means of hydrofluoric acid, and finally recovered in pure form from the acid solution according to conventional procedures.

As regards the peptide sequences of smallest size bearing an epitope or immunogenic determinant, and more particularly those which are readily accessible by chemical synthesis, it may be required, in order to increase their <u>in vivo</u> immunogenic character, to couple or "conjugate" them covalently to a physiologically acceptable and non-toxic carrier molecule. By way of examples of carrier molecules or macromolecular supports which can be used for making the conjugates according to

the invention can be mentioned natural proteins, such as tetanic toxoid, ovalbumin, serum-albumins, hemocyanins, etc. Synthetic macromolecular carriers, for example polysines or poly(D-L-alanine)-poly(L-lysine)s, can be used too. Other types of macromolecular carriers that can be used, which generally have molecular weights higher than 20,000, are known from the literature. The conjugates can be synthesized by known processes such as are described by Frantz and Robertson in "Infect. and Immunity", 33, 193-198 (1981) and by P.E. Kauffman in "Applied and Environmental Microbiology", October 1981 Vol. 42, No. 4, pp. 611-614. For instance, the following coupling agents can be used : glutaric aldehyde, ethyl chloroformate, water-soluble carbodiimides such as(N-ethyl-N'(3-dimethylamino-propyl) carbodiimide, HCl), diisocyanates, bis-diazobenzidine, di- and trichloro-s-triazines, cyanogen bromides and benzaquinone, as well as the coupling agents mentioned in "Scand. J. Immunol.", 1978, vol. 8, pp. 7-23 (Avrameas, Ternynck, Guesdon).

Any coupling process can be used for bonding one or several reactive groups of the peptide, on the one hand, and one or several reactive groups of the carrier, on the other hand. Again coupling is advantageously achieved between carboxyl and amine groups carried by the peptide and the carrier or vice-versa in the presence of a coupling agent of the type used in protein synthesis, e.g., 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide, N-hydroxybenzotriazole, etc. Coupling between amine groups respectively borne by the peptide and the carrier can also be made with glutaraldehyde, for instance, according to the method described by BOQUET, P. et al. (1982) Molec. Immunol., 19, 1441-1549, when the carrier is hemocyanin.

The immunogenicity of epitope-bearing-peptides

can also be reinforced by oligomerisation thereof, for example in the presence of glutaraldehyde or any other suitable coupling agent. In particular, the invention relates to the water soluble immunogenic oligomers thus obtained, comprising particularly from 2 to 10 monomer units.

The glycoproteins, proteins and other polypeptides (generally designated hereinafter as "antigens" of this invention) whether obtained by methods, such as are disclosed in the earlier patent applications referred to above, in a purified state from LAV_{MAL} virus preparations or - as concerns more particularly the peptides - by chemical synthesis, are useful in processes for the detection of the presence of anti-LAV antibodies in biological media, particularly biological fluids such as sera from man or animal, particularly with a view of possibly diagnosing LAS or AIDS.

Particularly the invention relates to an _in vitro_ process of diagnosis making use of an envelope glycoprotein or of a polypeptide bearing an epitope of this glycoprotein of LAV_{MAL} for the detection of anti-LAV antibodies in the serums of persons who carry them. Other polypeptides - particular those carrying an epitope of a core protein - can be used too.

A preferred embodiment of the process of the invention comprises :

- depositing a predetermined amount of one or several of said antigens in the cups of a titration microplate ;
- introducing increasing dilutions of the biological fluid, to be diagnosed (e.g., blood serum, spinal fluid, lymphatic fluid, and cephalo-rachidian fluid), into these cups ;
- incubating the microplate ;
- washing carefully the microplate with an appropriate buffer ;

- adding into the cups specific labelled antibodies directed against blood immunoglobulins and
- detecting the antigen-antibody-complex formed, which is then indicative of the presence of LAV antibodies in the biological fluid.

Advantageously the labelling of the anti-immunoglobulin antibodies is achieved by an enzyme selected from among those which are capable of hydrolysing a substrate, which substrate undergoes a modification of its radiation-absorption, at least within a predetermined band of wavelenghts. The detection of the substrate, preferably comparatively with respect to a control, then provides a measurement of the potential risks, or of the effective presence, of the disease.

Thus, preferred methods of immuno-enzymatic and also immunofluorescent detections, in particular according to the ELISA technique, are provided. Titrations may be determinations by immunofluorescence or direct or indirect immuno-enzymatic determinations. Quantitative titrations of antibodies on the serums studied can be made.

The invention also relates to the diagnostic kits themselves for the in vitro detection of antibodies against the LAV virus, which kits comprise any of the polypeptides identified herein and all the biological and chemical reagents, as well as equipment, necessary for peforming diagnostic assays. Preferred kits comprise all reagents required for carrying out ELISA assays. Thus preferred kits will include, in addition to any of said polypeptides, suitable buffers and anti-human immunoglobulins, which anti-human immunoglobulins are labelled either by an immunofluorescent molecule or by an enzyme. In the last instance, preferred kits also comprise a substrate hydrolysable by the enzyme and providing a signal, particularly modified absorption of a

radiation, at least in a determined wavelength, which signal is then indicative of the presence of antibody in the biological fluid to be assayed with said kit.

It can of course be of advantage to use several proteins or polypeptides not only of LAV_{MAL}, but also of LAV_{ELI} together with homologous proteins or polypeptides of earlier described viruses, such as LAV_{BRU}, HTLV-3, ARV 2, etc.

The invention also relates to vaccine compositions whose active principle is to be constituted by any of the antigens, i.e., the hereinabove disclosed polypeptides of LAV_{MAL}, particularly the purified gp110 or immunogenic fragments thereof, fusion polypeptides or oligopeptides in association with a suitable pharmaceutically or physiologically acceptable carrier. A first type of preferred active principle is the gp110 immunogen of said immunogens. Other preferred active principles to be considered in that fields consist of the peptides containing less than 250 amino acid units, preferably less than 150, particularly from 5 to 150 amino acid residues, as deducible for the complete genome of LAV_{MAL} and even more preferably those peptides which contain one or more groups selected from Asn-X-Thr and Asn-X-Ser as defined above. Preferred peptides for use in the production of vaccinating principles are peptides (a) to (f) as defined above. By way of example, there may be mentioned that suitable dosages of the vaccine compositions are those which are effective to elicit antibodies <u>in vivo</u>, in the host, particularly a human host. Suitable doses range from 10 to 500 micrograms of polypeptide, protein or glycoprotein per kg, for instance 50 to 100 micrograms per kg.

The different peptides according to this invention can also be used themselves for the production of antibodies, preferably monoclonal antibodies specific

for the respective different peptides. For the production of hybridomas secreting said monoclonal antibodies, conventional production and screening methods can be used. These monoclonal antibodies, which themselves are part of the invention, provide very useful tools for the identification and even determination of relative proportions of the different polypeptides or proteins in biological samples, particularly human samples containing LAV or related viruses.

The invention further relates to the hosts (procaryotic or eucaryotic cells) which are transformed by the above mentioned recombinants and which are capable of expressing said DNA fragments.

Finally the invention also concerns vectors for transforming eucaryotic cells of human origin, particularly lymphocytes, the polymerase of which are capable of recognizing the LTRs of LAV. Particularly said vectors are characterized by the presence of a LAV LTR therein, said LTR being then active as a promoter enabling the efficient transcription and translation in a suitable host of a DNA insert coding for a determined protein placed under its controls.

Needless to say, the invention extends to all variants of genomes and corresponding DNA fragments (ORFs) having substantially equivalent properties, all of said genomes belonging to retroviruses which can be considered as equivalents of LAV_{MAL}. It must be understood that the claims which follow are also intended to cover all equivalents of the products (glycoproteins, polypeptides, DNAs, etc.) whereby an equivalent is a product, e.g., a polypeptide, which may distinguish from a product defined in any of said claims, say through one or several amino acids, while still having substantially the same immunological or immunogenic properties. A similar rule of equivalency shall apply to the DNAs, it

being understood that the rule of equivalency will then be tied to the rule of equivalency pertaining to the polypeptides which they encode.

It will also be understood that all the literature referred to hereinbefore and hereinafter and all patent applications and patents not specifically identified herein but which form counterparts of those specifically designated herein, must be considered as incorporated herein by reference.

It should further be mentioned that the invention further relates to immunogenic compositions that contain preferably one or more of the polypeptides, which are specifically identified above and which have the amino acid sequences of LAV_{MAL} that have been identified, or peptidic sequences corresponding to previously defined LAV proteins. In this respect, the invention relates more particularly to the particular polypeptides which have the sequences corresponding more specifically to the LAV_{BRU} sequences which have been referred to earlier, i.e., the sequences extending between the following first and last amino acids, of the LAV_{BRU} proteins themselves, i.e., the polypeptides having sequences contained in the LAV_{BRU} OMP or LAV_{BRU} TMP or sequences extending over both, particularly those extending from between the following positions of the amino acids included in the <u>env</u> open reading frame of the LAV_{BRU} genome,

1-530

34-530

and more preferably

531-877, particularly 680-700,

37-130

211-289

488-530

490-620.

These different sequences can be used for any of the above defined purposes and in any of the compositions which have been disclosed.

Finally the invention also relates to the different antibodies which can be formed specifically against the different peptides which have been disclosed herein, particularly to the monoclonal antibodies which recognize them specifically. The corresponding hybridomas which can be formed starting from spleen cells previously immunized with such peptides which are fused with appropriate myeloma cells and selected according to standard procedures also form part of the invention.

Phage λ clone E-H12 derived from LAV_{ELI} infected cells has been deposited at the CNCM under No. I-550 on May 9, 1986. Phage clone M-H11 derived from LAV_{MAL} infected cells has been deposited at the CNCM under No. I-551 on May 9, 1986.

REFERENCES

Alizon, M., Sonigo, P., Barré-Sinoussi, F., Chermann, J.C., Tiollais, P., Montagnier, L. & Wain-Hobson, S. (1984). Molecular cloning of lymphadenopathy-associated virus. Nature 312, 757-760.

Allan, J.S., Coligan, J.E., Barin, F., McLane, M.F., Sodroski, J.G., Rosen, C.A., Haseltine, W.A., Lee, T.H., & Essex, M. (1985a). Major glycoprotein antigens that induce antibodies in AIDS patients. Science 228, 1091-1094.

Allan, J.S., Coligan, J.E., Lee, T.H., McLane, M.F., Kanki, P.J., Groopman, J.E., & Essex, M. (1985b). A new HTLV-III/LAV antigen detected by antibodies from AIDS patients. Science 230, 810-813.

Arya, S.K., Guo, C., Josephs, S.F., & Wong-Staal, F. (1985). Trans-activator gene of human T-lymphotropic virus type III (HTLV-III). Science 229, 69-73.

Bailey, A.C., Downing, R.G., Cheinsong-Popov, R., Tedder, R.C., Dalgleish, A.G., & Weiss, R.A.(1985). HTLV-III serology distinguishes atypical and endemic Kaposi's sarcoma in Africa. Lancet I, 359-361.

Barin, F., M'Boup, S., Denis, F., Kanki, P., Allan, J.S., Lee, T.M., & Essex, M. (1985). Serological evidence for virus related to simian T lymphotropic retrovirus in residents of West Africa. Lancet II, 1387-1389.

Barré-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Brun-Vézinet, F., Rouzioux, C., Rozenbaum, W. & Montagnier, L. (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk of acquired immune deficiency syndrome (AIDS). Science 220, 868-870.

Been, S., Rutledge, R., Folks, T., Gold, J., Baker, L. McCormick, J. Feorino, P., Piot, P., Quinn T. & Martin, M. (1985). Genomic heterogeneity of AIDS retroviral

isolates from North America and Zaire. Science 230, 949-951.

Borst, P., & Cross, G.A.M. (1982). Molecular basis for trypanosome antigenic variation. Cell 29, 291-303.

Brun-Vésinet, F., Rouzioux, C., Montagnier, L., Chamaret, S., Gruest, J., Barré-Sinoussi, F., Geroldi, D., Chermann, J.C., McCormick, J. Mitchell, S., Piot, P., Taelmann, H. Minlangu, K.B., Wobin, O., Mbendi, N. Mazebo, P., Kalambayi, K. Bridts, C., Desmyter, J., Feinsod, F., & Quinn T.C. (1984). Prevelance of antibodies to lymphadenopathy-associated virus in African patients with AIDS. Science 226, 453-456.

Chang, N.T., Chanda, P.K., Barone, A.D., McKinney, S., Rhodes, D.P., Tam, S.H., Shearman, C.W., Huang, J. & Chang, T.W. (1985). Expression in Escheridia coli of open reading frame gene segments of HTLV-III. Science 228, 93-96.

Chiu, I.M., Yaniv, A., Dahlberg, J.E., Gazit, A., Skuntz, S.F., Tronick, S.R. & Aaronson, S.A. (1985). Nucleotide sequence evidence for relationship of AIDS retrovirus to lentiviruses. Nature 317, 366-368.

Clark, S.P., & Mak, T.W., (1984). Fluidity of a retrovirus genome. J. Virol. 50, 759-765.

Clavel, F., Klatzmann, D., & Montagnier, L., (1985). Deficient neutralizing capacity of sera from patients with AIDS or related syndromes. Lancet I, 879-880.

Clavel, F., Brun-Vézinet, F., Guetard, D., Chamaret, S., Laurent, A., Rouzioux, C., Rey, M., Katlama, C., Rey, F., Champelinaud, J.L., Nina, J, S., Mansinho, K., Santos-Ferreira, M.O., Klatzmann, D., & Montagnier, L. (1986). LAV type II: a second retrovirus associated with AIDS in West-Africa. C.R.Acad.Sci.Paris 302, 485-488.

Clements, J.E., Narayan, O., Griffin, D.E. and Johnson, R.T. (1980). genomic changes associated with antigenic variation of visna virus during persistent infection.

Proc. Natl. Acad. Sci. USA 77,4454-4458.

Clumeck, N., Sonnet, J., Taelman, M., Mascart-Lemone, F., De Bruyère, M., Van de Perre, P., Dasnoy, J., Marcelis, L., Lamy, M., Jonas, C., Eyckmans, L., Noel, H., Vanhaeverbeek, M. & Butzler, J.P. (1984). Acquired immune deficiency syndrome in African patients. N. Engl. J. Med., 10, 492-497.

Dalgleish, A.G., Beverley, P.C., Clapham P.R., Crawford, D.H., Greaves, M.F. & Weiss, R.A. (1984). The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. Nature 312, 763-767.

Darlix, J.L. (1986) Control of Rous sarcoma virus RNA translation and packaging by the 5' and 3' untranslated sequences. J.Mol.Biol., in the press.

DeLarco, J. & Todaro, G.J. (1976). Membrane receptors of murine leukemia viruses: characterization using the purified viral envelope glycoprotein, gp71. Cell 8, 365-371.

DiMarzoVeronese, F., DeVico, A.L., Copeland, T.D., Oroszlan, S., Gallo, R.C., & Sarngadharan, M.G. (1985). Characterization of gp 41 as the transmembrane protein coded by the HTLV-III/LAV envelope gene. Science 229, 1403-1405.

Ellrodt, A., Barré-Sinoussi, F., Le Bras, P., Nugeyre, M.T., Brun-Vézinet, F., Rouzioux, C., Segond, P., Caquet, R., Montagnier, L. & Chermann, J.C. (1984). Isolation of human T-lymphotropic retrovirus (LAV) from Zairan married couple, one with AIDS, one with prodromes. Lancet I, 1383-1385.

Hahn, B.H., Shaw, G.M., Arya, S.U., Popovic, M., Gallo, R.C., & Wong-Staal, F. (1984). Molecular cloning and characterizaion of the HTLV-III virus associated with AIDS. Nature 312, 166-169.

Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., & Van de Pol, S. (1982). Rapid evolution of

RNA genomes. Science 215, 1577-1585.

Kan, N.C., Franchini, G., Wong-Staal, F., Dubois, G.C., Robey, W.G., Lautenberger, J.A., & Papas, T.S. (1986). Identification of HTLV-III/LAV sor gene product and detection of antibodies in human sera. Science 231, 1553-1555.

Kanki, P.J., Alroy, J. & Essex, M. (1985). Isolation of T-lymphotropic retroviruses from wild-caught African Green Monkeys. Science 230, 951-954.

Kanki, P.J., Barin, F., M'Boup, S., Allan, J.S., Romet-Lemonne, J.L., Markink, R., McLane, M.F., Lee, T.H., Arbeille, B., Denis, F. & Essex, M. (1986). New human T-lymphotropic retrovirus related to simian T-lymphotropic virus type III (STLV-III$_{AGM}$). Science, 232, 238-243.

Klatzmann, D., Champagne, E., Chamaret, S., Gruest, J., Guetard, D., Hercend, T., Gluckman, J.C., & Montagnier, L. (1984). T-lymphocyte T4 molecule behave as the receptor for human retrovirus LAV. Nature 312, 767-768.

Kyte, J. & Doolittle, R., (1982). A simple methof for displaying the hydropathic character of a protein. J.Mol.Biol. 157, 105-132.

Koch, W., Hunsmann, G. & Friedrich, R. (1983). Nucleotide sequence of the envelope gene of Friend murine leukemia virus. J. Virol., 45, 1-9.

Lee, T.H., Coligan, J.E. Allan, J.S., McLane, M.F., Groopman, J.E. & Essex, M. (1986). A new HTLV III/LAV protein encoded by a gene found in cytopathic retroviruses. Science 231, 1546-1549.

Loenec, W.A.M. & Brammar, W.J. (1980). A bacteriophage lambda vector for cloning large DNA fragments made with several restriction enzymes. Gene 10, 249-259.

Lutley, R., Petursson, G., Palsson, P.A., Georgsson, G., Klein, J., & Nathanson, N. (1983). Antigenic drift in visna: virus variation during longterm infection of

icelandic sheep. J. Gen. Virol. 64, 1433-1440.

MacDougal, J.S., Kennedy, M.S., Sligh, J.M., Cort, S.P., Mawle, A. & Nicholson, J.K.A. (1986). Binding of HTLV-III/LAV to T4$^+$ cells by a complex of the 110 k viral protein and the T4 molecule. Science 231, 382-385.

Messing, J. and Viera, J. (1982). A new pair of M13 vectors for selecting either DNA strand of double digest restriction fragments. Gene 19, 269-276.

Montagnier, L. (1985). Lymphadenopathy-associated virus:from molecular biology to pathogenicity. Ann.Inter.Med. 103, 689-693.

Montagnier, L., Clavel, F., Krust, B., Chamaret, S., Rey, F., Barré-Sinoussi, F. & Chermann, J.C. (1985). Identification and antigenicity of the major envelope glycoprotein of lymphadenopathy-associated virus. Virology 144, 283-289.

Montelaro, R.C, Parekh, B., Orrego, A. & Issel, C.J. (1984). Antigenic variation during persistent infection by equine infectious anemia virus, a retrovirus. J.Biol.Chem., 250, 10539-10544.

Muesing, M.A., Smith, D.M., Cabradilla, C.D., Benton, C.V., Lasky, L.A. & Capon, D.J. (1985). Nucleic acid structure and expression of the human AIDS/lymphadenopathy retroviruses. Nature 313, 450-458.

Norman, C. (1985). Politics and science clash on African AIDS. Science 230, 1140-1142.

Piot, P., Quinn, T.C., Taelman, H., Feinsod, F.M., Minlangu, K.B., Wobin, O., Mbendi, N., Mazebo, P., Ndongi, K., Stevens, W., Kalambayi, K., Mitchell, S., Bridts, C. & McCormick, J.B. (1984). Acquired immunodeficiency syndrome in

Power, M.D., Marx, P.A., Bryant, M.L., Gardner, M.B., Barr, P.J. & Luciw, P.A. (1986). Nucleotide sequence of SRV-1, a type D simian acquired immune deficiency syndrome retrovirus. Science 231, 1567-1572.

Rabson, A.B. & Martin, M.A. (1985). Molecular organization of the AIDS retrovirus. Cell 40, 477-480.

Ratner, L., Haseltine, W., Patarca, R., Livak, K.J., Starcich, B., Josephs, S.F., Doran, E.R., Rafalski, A., Whitchorn, E.A., Baumeister, K., Ivanoff, L., Petteway, S.R., Pearson, M.L., Lautenbergen, J.A., Papas, T.S., Ghrayeb, J., Chang, N.T., Gallo, R.C. & Wong-Staal, F. (1985). Complete nucleotide sequence of the AIDS virus, HTLV-III. Nature 313, 277-284.

Robinson, P.J.G., Hunsmann, G., Schneider, J. & Schirrmacher, V. (1980). Possible cell surface receptor for Friend murine leukemia virus is isolated with viral envelop glycoprotein complexes. J.Virol., 36, 291-294.

Rosen, C.A., Sodroski, J.G. & Haseltine, W.A. (1985). The location of cisacting regulatory sequences in the human T cell lymphotropic virus type III (HTLV-III/LAV) long terminal repeat. Cell 41, 813-823.

Sanchez-Pescador, R. Power, M.D., Barr, P.J., Steimer, K.S., Stemfeieb, M.M., Brown-Shimer, S.L., Gee, W.W., Bernard, A., Randolph, A., Levy, J.A., Dina, D. & Luciw, P.A., (1985). Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2). Science 227, 484-492.

Sanger, F., Nicklen, S. & Coulsen, A.R. (1977). DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463-5467.

Scott, J.V., Stowring, L., Haase, A.T., Narayan, O. & Vigne, R. (1979). Antigenic variation in visna virus. Cell 18, 321-327.

Shimotohno, K., & Temin, H. (1982). Spontaneous variation and synthesis in the U3 region of the long terminal repeat of an avian retrovirus. J. Virol. 41, 163-171.

Sodroski, J., Patarca, R., Rosen C., Wong-Staal, F. & Haseltine, W. (1985). Location of the trans-activating

region of the genome of human T-cell lymphotropic virus type II. Science 229, 74-77.

Sonigo, P., Alizon, M., Staskus, K., Klatzmann, D., Cole, S., Danos, O., Retzel, E., Tiollais, P., Haase, A. & Wain-Hobson, S. (1985). Nucleotide sequence of the visna lentivirus: Relationship to the AIDS virus. Cell 42, 369-382.

Sonigo, P., Barker, C., Hunter, E. & Wain-Hobson S. (1986). Nucleotide sequence of Mason-Pfizer Monkey virus: an immunosuppressive D-type retrovirus. Cell, in the press.

Steinhauer, D.A., & Holland, J.H. (1986). Direct method for quantitation of extreme polymerase error frequencies at selected single base in viral RNA. J. Virol. 57, 219-228.

Thormar, H., Barshatsky, M.R., Arnesen, K., & Kozlowski, P.B. (1983). The emergence of antigenic variants is a rare event in long-term visna virus invention <u>in vivo</u>. J. Gen. Virol. 64, 1427-1432.

Van de Perre, P., Rouvroy, D., Lepage, P., Bogaerts, J., Kestelyn, P., Kayihigi, J., Hekker, A.C., Butzler, J.P. & Clumeck, N. (1984). Acquired immunodeficiency syndrome in Rwanda. Lancet II, 62-65.

Varmus, H. & Swanstrom, R. (1984). Replication of retroviruses. In Molecular biology of the tumor viruses/ RNA tumor viruses. R. Weiss, N. Teich, H. Varmus, J. Coffin, eds. (Cold Spring Harbor Laboratory, New York), vol. 1, pp. 369-512.

Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S., & Alizon, M. (1985). Nucleotide sequence of the AIDS virus, LAV. Cell 40, 9-17.

Webster, R.G., Laver, W.G., Air, G.M. & Schild, G.C. (1982). Molecular mechanisms of variation in influenza viruses. Nature 296, 115-121.

Weiss, R.A. (1984). Human T-cell retroviruses. In

Molecular biology of the tumor viruses: RNA viruses. R. Weiss, N. Teich, H. Varmus, J. Coffin, eds. (Cold Spring Harbor Laboratory, New York), vol II: supplement, pp. 405-485.

Weiss, R.A., Clapham, P.R., Cheinson-Popov, R. Dalgleish, A.G., Carne, C.A., Weller, I.A.D. & Tedder, R.C. (1985). Neutralization of human T-lymphotropic virus type III by sera of AIDS and AIDS-risk patients, Nature, 316, 69-72.

Wilburg, W.J., & Lipman, D.J. (1983). Rapid similarity searches of nucleic acid and protein data banks. Proc. Natl. Acad. Sci. USA 80, 726-730.

Wu, T.T., & Kabat, E.A. (1970). An analysis of the sequences of the variable regions of Bence-Jones proteins and myeloma light chains and their implications for antibody complementarity. J.Exp.Med. 132, 211-250.

Claims:

1. The virus LAV_{MAL} comprising RNA corresponding to the cDNA of figs. 7A-7I.

2. The cDNA of figs. 7A-7I.

3. A DNA recombinant comprising the cDNA of claim 2.

4. A probe containing a nucleic acid sequence hybridizable with RNA of said LAV_{MAL} virus of claim 1.

5. A method for identifying the presence in a host tissue of LAV virus which comprises hybridizing RNA obtained from said tissue with said probe of claim 4.

6. The method of claim 5, wherein said probe can hybridize with RNA from said LAV_{MAL} virus to identify said LAV_{MAL} virus.

7. A peptide or fragment thereof whose amino acid sequence is encoded by an open reading frame of a cDNA sequence of the LAV_{MAL} virus of claim 1.

8. The peptide of claim 7 encoded by a cDNA sequence from amino-acyl residue 37 to amino-acyl residue 130, or from amino-acyl residue 211 to amino-acyl residue 289, or from amino-acyl residue 488 to amino-acyl residue 530 of figs. 3A-3F and 7A-7I.

9. The peptide of claim 7 encoded by a cDNA sequence from amino-acyl residue 490 to amino-acyl residue 620 or from amino-acyl residue 680 to amino-acyl residue 700 of figs. 3A-3F and 7A-7I.

10. The peptide of claim 7 which comprises a protein or glycoprotein whose amino acid sequence is encoded by all or part of one of the following cDNA sequences of figs. 3A-3F and 7A-7I:

OMP or gp110 proteins, including precursors: 1 to 530;

OMP or gp110 without precursor: 34-530; and

TMP or gp41 protein: 531-877.

11. The peptide of claim 10 encoded by all

or part of one of the following cDNA sequences of figs. 3A-3F and 7A-7I: 37-130, 211-289, 488-530, 490-620 or 680-700.

12. A method for the in vitro detection of the presence of an antibody directed against a LAV virus in a human body fluid, which comprises: contacting said body fluid with an antigen obtained from said virus LAV_{MAL} of claim 1, said antigen consisting of a peptide or a fragment thereof whose amino acid sequence is encoded by an open reading frame of a cDNA sequence of figs. 7A-7I; and then detecting the immunological reaction between said antigen and said antibody.

13. The method of claim 12 wherein said antigen detects said LAV_{MAL} virus of claim 1.

14. The method of claim 12 which comprises the steps of:

a) depositing a predetermined amount of said antigen into a cup of a titration microplate;

b) introducing increasing dilutions of said body fluid into said cup;

c) incubating said microplate;

d) washing the microplate with a buffer;

e) adding into said cup a labelled antibody directed against blood immunoglobulins; and then

f) determining whether an antigen-antibody-complex has formed in said cup which is indicative of the presence of a LAV antibody in said body fluid.

15. A diagnostic kit for the in vitro detection of antibodies against a LAV virus, which kit comprises: an antigen consisting of a peptide of claim 7.

16. The kit of claim 15 wherein the antigen consists of a peptide of said LAV_{MAL} virus of claim 1, encoded by the open reading frame of a cDNA sequence of said LAV_{MAL} virus.

17. An immunogenic composition comprising: an antigen of the LAV_{MAL} virus of claim 1 or an immunogenic peptide or fragment thereof encoded by RNA of said virus; and a physiologically acceptable carrier.

18. The immunogenic composition of claim 17 wherein said peptide is the gp110 envelope glycoprotein or a fragment thereof.

19. The immunogenic composition of claim 17 wherein the peptide comprises a protein or glycoprotein hwose amino acid sequence is encoded by all or part of one of the following cDNA sequences of figs 3a-3F and 7A-7I:

OMP or gp110 proteins, including precursors: 1 to 530;

OMP or gp110 without precursor: 34-530; and

TMP or gp41: 531-877.

20. The composition of claim 19 wherein the protein or glycoprotein is encoded by all or part of one of the following cDNA sequences of Figs. 3A-3F and 7A-7I: 37-130, 211-289, 488-530, 490-620 or 680-700.

21. An antibody formed against a peptide of claim 7.

22. A cell transformed with a DNA recombinant of claim 3.

ABSTRACT

A variant of a LAV virus, designated LAV_{MAL} and capable of causing AIDS. The cDNA and antigens of the LAV_{MAL} virus can be used for the diagnosis of AIDS and pre-AIDS.

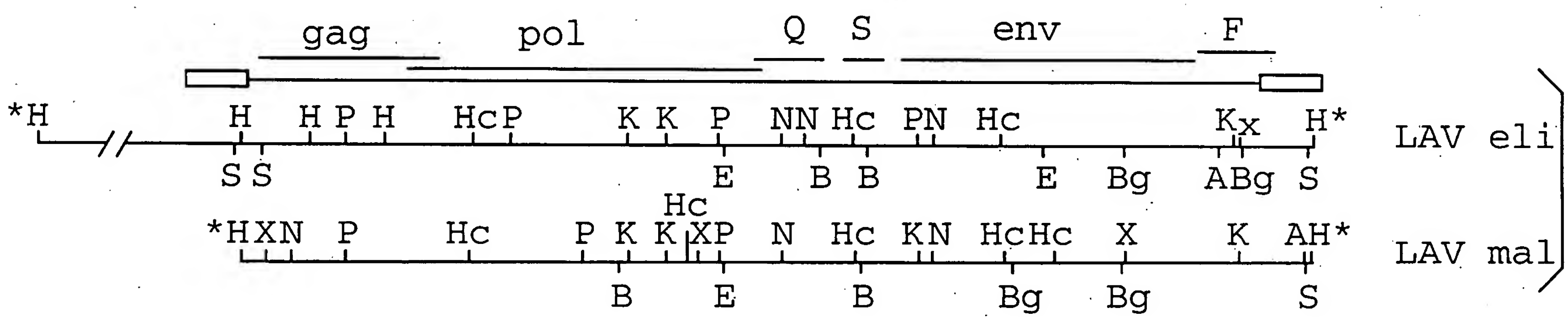


FIG. 1A

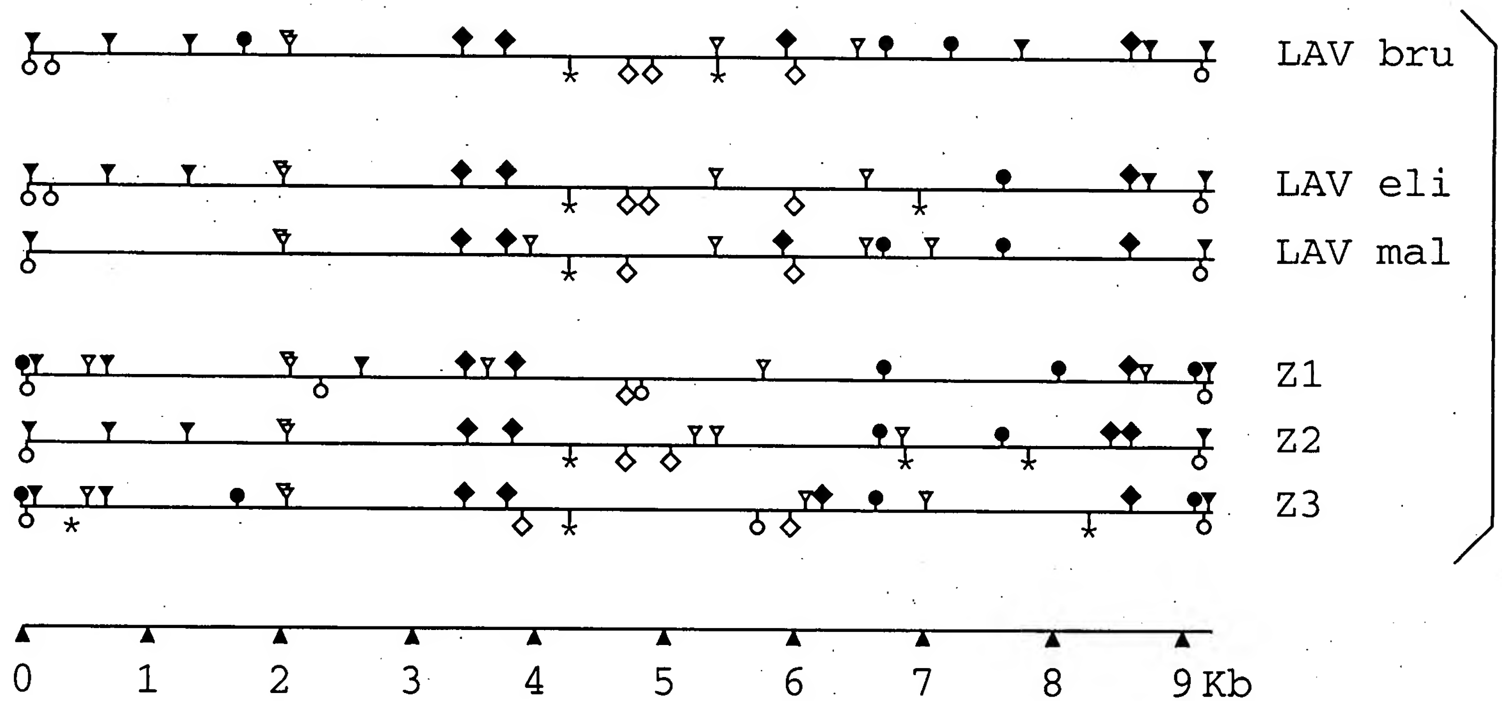


FIG. 1B

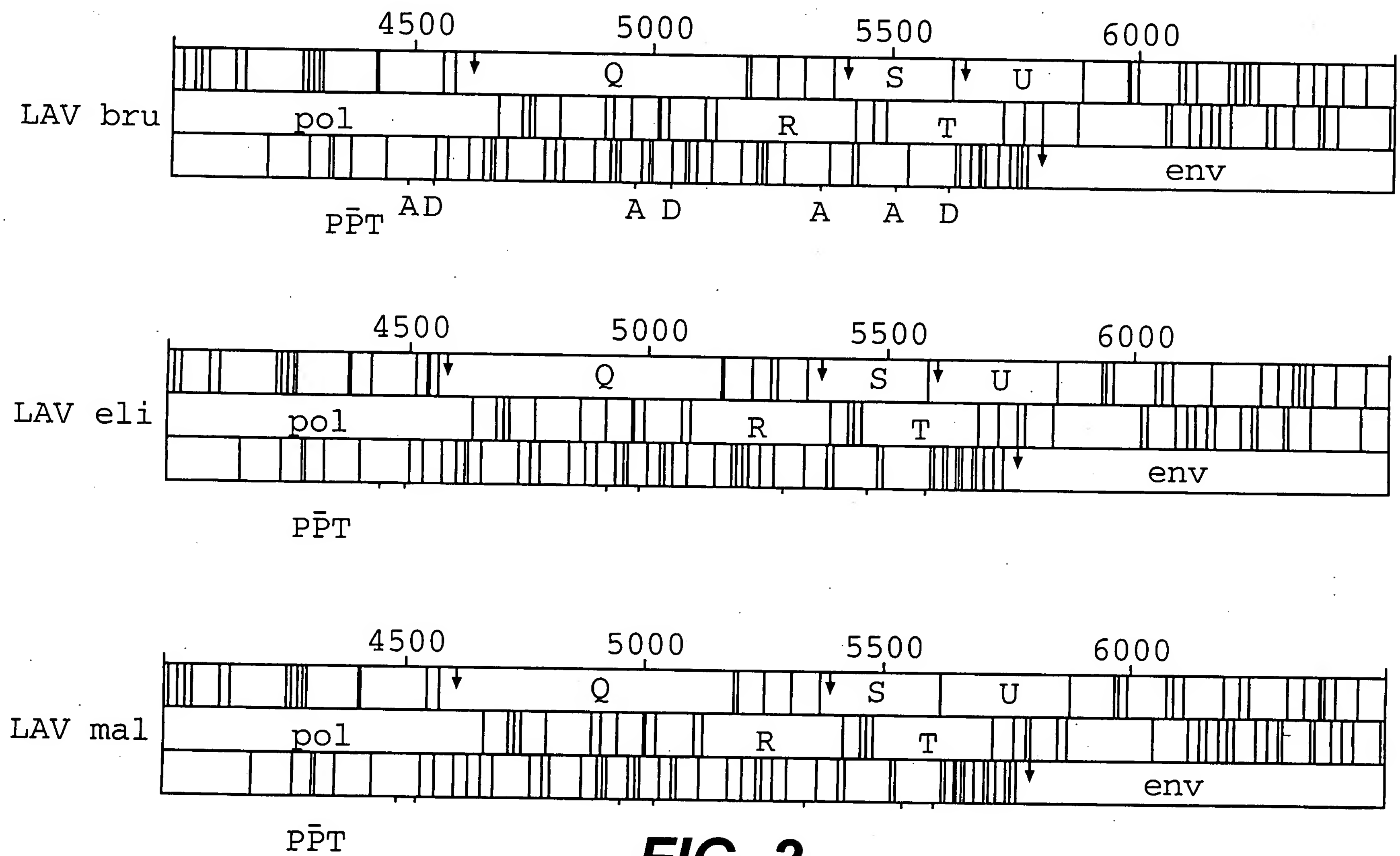

4500
5000
5500
6000
LAV bru
Q
S
U
pol
R
T
env
PP̄T
AD
A D
A
A D
LAV eli
PP̄T
LAV mal

FIG. 2

GAG

```
                  10         20         30         40         50         60         70         80
LAV BRU   MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL LETSEGCRQI LGQLQPSLQT GSEELRSLYN
ARV 2                    K
LAV MAL               K  A               R    L            L        C   Q   ME   ST K      IK
LAV ELI               K  K               R               Y L            K   I     AI    T
                                                                        ↓p25
                  90        100        110        120        130        140        150        160
LAV BRU   TVATLYCVHQ RIEIKDTKEA LDKIEEEQNK SKKKAQQAAA ------DTGH SSQVSQNYPI VQNIQGQMVH QAISPRTLNA
ARV 2                  DV        E                    ----AAG  N               L
LAV MAL                DV            I     RQ T       AQQAAAA KN   S           A    I
LAV ELI            K G DV        E M                  ------   N N             L

                 170        180        190        200        210        220        230        240
LAV BRU   WVKVVEEKAF SPEVIPMFSA LSCGATPQDL NTMLNTVGGH QAAMQMLKET INEEAAEWDK VHPVHAGPIA PGQMREPRGS
ARV 2
LAV MAL       I                             M   I             D        D             P
LAV ELI       I                                                             L

                 250        260        270        280        290        300        310        320
LAV BRU   DIAGTTSTLQ EQIGWMTNNP PIPVGEIYKR WIILGLNKIV RMYSPTSILD IRQGPKEPFR DYVDRFYKTL RAEQASQEVK
ARV 2                                                                                         D
LAV MAL                     S        D                     V                      F         T
LAV ELI                 A   S                 V            V                                  D
```

FIG. 3A-1

```
                                                                              |p13
                 330        340        350        360        370        380    ↓  390        400
LAV BRU   NWMTETLLVQ NANPDCKTIL KALGPAATLE EMMTACQGVG GPGHKARVLA EAMSQVTNS- ATIMMQRGNF RNQRKIVKCF
ARV 2                                                                  P-  N              T
LAV MAL                                G                 S          A   T  A          KG - RI
LAV ELI                                Q                 S          A   V T A         KGP   I

                 410        420        430        440        450        460        470        480
LAV BRU   NCGKEGHIAR NCRAPRKKGC WKCGKEGHQM KDCTERQANF LGKIWPSYKG RPGNFLQSRP EPTAPPFLQS RPEPTAPPEE
ARV 2              K             R   R                                             ---- --------
LAV MAL          L                                            H                    ---- --------A
LAV ELI            K             R        L              R    H                    ---- --------A

                 490        500        510
LAV BRU   SFRSGVETTT PSQKQEPIDK ELYPLTSLRS LFGNDPSSQ
ARV 2        F E K
LAV MAL     GF E IK-        QK        A  K       QL
LAV ELI     GF E I -        QK           K        L
```

FIG. 3A-2

```
CENTRAL REGION:  Q
                  10         20         30         40         50         60         70         80
LAV BRU   MENRWQVMIV WQVDRMRIRT WKSLVKHHMY VSGKARGWFY RHHYESPHPR ISSEVHIPLG DARLVITTYW GLHTGERDWH
ARV 2                                      I K  K           T    V            K               E
LAV MAL                          H           K  KN          R  K V               VR      Q   K
LAV ELI                      K               K  NR             K            E     K           E

                  90        100        110        120        130        140        150        160
LAV BRU   LGQGVSIEWR KKRYSTQVDP ELADQLIHLY YFDCFSDSAI RKALLGHIVS PRCEYQAGHN KVGSLQYLAL AALITPKKIK
ARV 2          A       K        G        H       E    KN I  YR                                 T
LAV MAL     H        Q      L   D                E     Q I          D                  T   A   TR
LAV ELI               R         G       M        E       I  D                          T   A  Q

                 170        180        190
LAV BRU   PPLPSVTKLT EDRWNKPQKT KGHRGSHTMN GH
ARV 2           K
LAV MAL         R           Q
LAV ELI         R           Q  R
```

FIG. 3B-1

```
                R
                   10         20         30         40         50         60         70         80
LAV BRU   MEQAPEDQGP QREPHNEWTL ELLEELKNEA VRHFPRIWLH GLGQHIYETY GDTWAGVEAI IRILQQLLFI HFRIGCRHSR
ARV 2                    Y             R          P    S   Y                                  Q
LAV MAL        A                       Q               S              E        S              Q
LAV ELI        A         Y   A         S               S              V                       Q

                   90
LAV BRU   IGVTQQRRAR -NGASRS
ARV 2       II       R
LAV MAL     I R      -  S
LAV ELI     IIR      -  S

              S (tat)

                   10         20         30         40         50         60         70
LAV BRU   MEPVDPRLEP WKHPGSQPKT ACTTCYCKKC CFHCQVCFTT KALGISYGRK KRRQRRRPPQ GSQTHQVSLS KQ
ARV 2           N            R    NN            YA   R  G                 A   D     A
LAV MAL    D    N     N      R  P NK          Y   M  I  G                      N A  DP P E
LAV ELI    D    N     N      R  P NK H        Y  P   LN G               G      G A   PIP
```

FIG. 3B-2

POL

```
                 10         20         30         40         50         60         70         80
LAV BRU  FFREDLAFLQ GKAREFSSEQ TRANSPTFSS EQTRANSPTR RELQVWGRDN NSLSEAGADR QGTVSFNFPQ ITLWQRPLVT
ARV 2                                --- ---------          GE
LAV MAL      N   P        P          --- ---------S     R   G - KT   T  E   I    S           V
LAV ELI      N   P     G L PK        --- ---------S     R     -  P  KT  E                      A

                 90        100        110        120        130        140        150        160
LAV BRU  IKIGGQLKEA LLDTGADDTV LEEMSLPGRW KPKMIGGIGG FIKVRQYDQI LIEICGHKAI GTVLVGPTPV NIIGRNLLTQ
ARV 2     R                         N   K                        PV
LAV MAL  VRV                       IN   K                              K      I              M
LAV ELI                             N   K                        P     Q

                170        180        190        200        210        220        230        240
LAV BRU  IGCTLNFPIS PIETVPVKLK PGMDGPKVKQ WPLTEEKIKA LVEICTEMEK EGKISKIGPE NPYNTPVFAI KKKDSTKWRK
ARV 2
LAV MAL                              R                T    KD       L
LAV ELI                                               T     D        R           I

                250        260        270        280        290        300        310        320
LAV BRU  LVDFRELNKR TQDFWEVQLG IPHPAGLKKK KSVTVLDVGD AYFSVPLDED FRKYTAFTIP SINNETPGIR YQYNVLPQGW
ARV 2                                                       K
LAV MAL    N
LAV ELI                                                                 S
```

FIG. 3C-1

```
                330        340        350        360        370        380        390        400
LAV BRU  KGSPAIFQSS MTKILEPFRK QNPDIVIYQY MDDLYVGSDL EIGQHRTKIE ELRQHLLRWG LTTPDKKHQK EPPFLWMGYE
ARV 2                                                                       F
LAV MAL                      T K  E                                 E   K   F
LAV ELI                           EM                             K  E       F R

                410        420        430        440        450        460        470        480
LAV BRU  LHPDKWTVQP IVLPEKDSWT VNDIQKLVGK LNWASQIYPG IKVRQLCKLL RGTKALTEVI PLTEEAELEL AENREILKEP
ARV 2                M                            A     K
LAV MAL              Q  D E                             K         A    DIV    A
LAV ELI           S  K    E         N  ER

                490        500        510        520        530        540        550        560
LAV BRU  VHGVYYDPSK DLIAEIQKQG QGQWTYQIYQ EPFKNLKTGK YARTRGAHTN DVKQLTEAVQ KITTESIVIW GKTPKFKLPI
ARV 2      E          V                                  M                   VS          I
LAV MAL                                     QY           IKS                  AQ             R
LAV ELI                         H                        M            A     R S         R    R
```

FIG. 3C-2

```
                570        580        590        600        610        620        630        640
LAV BRU  QKETWETWWT EYWQATWIPE WEFVNTPPLV KLWYQLEKEP IVGAETFYVD GAASRETKLG KAGYVTNRGR QKVVTLTDTT
ARV 2          A  M                                                N        D        SIA
LAV MAL        A                                 T                 N    K   D        S  E
LAV ELI           A                                   I            N        D        P

                650        660        670        680        690        700        710        720
LAV BRU  NQKTELQAIH LALQDSGLEV NIVTDSQYAL GIIQAQPDKS ESELVNQIIE QIIKKEKVYL AWVPAHKGIG GNEQVDKLVS
ARV 2                                                     S
LAV MAL                    S                            I          Q D     S
LAV ELI           N

                730        740        750        760        770        780        790        800
LAV BRU  AGIRKVLFLD GIDKAQDEHE KYHSNWRAMA SDFNLPPVVA KEIVASCDKC QLKGEAMHGQ VDCSPGIWQL DCTHLEGKVI
ARV 2             N       E
LAV MAL  S                E                      I                                            I
LAV ELI Q                 E       N                                                           I

                810        820        830        840        850        860        870        880
LAV BRU  LVAVHVASGY IEAEVIPAET GQETAYFLLK LAGRWPVKTI HTDNGSNFTS TTVKAACWWA GIKQEFGIPY NPQSQGVVES
ARV 2
LAV MAL  I                            I           VV            AA         N
LAV ELI                                           VV            AA
```

FIG. 3D-1

```
                890        900        910        920        930        940        950        960
LAV BRU  MNKELKKIIG QVRDQAEHLK TAVQMAVFIH NFKRKGGIGG YSAGERIVDI IATDIQTKEL QKQITKIQNF RVYYRDSRDP
ARV 2      N                                                                                KK
LAV MAL                E                                    I M                             N
LAV ELI                                       RR            I                  I

                970        980        990       1000       1010
LAV BRU  LWKGPAKLLW KGEGAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED
ARV 2
LAV MAL  I                                                    G G
LAV ELI  I                     K          V
```

FIG. 3D-2

ENV

```
                          SP                               OMP
                 10         20         30   ↓     40         50         60         70         80
LAV BRU  MRVK---EKY QHLWRWGWKW GTMLLGILMI CSATEKLWVT VYYGVPVWKE ATTTLFCASD AKAYDTEVHN VWATHACVPT
ARV 2     K  GTRRN             --- - L   M                                  R
LAV MAL     REIQRN   NW        --- -M    M   T    IA D                          S E       I
LAV ELI    ARGIERNC  NW K      --- -I        T     ADN                          S E   A   I

                 90        100        110        120        130        140        150        160
LAV BRU  DPNPQEVVLV NVTENFNMWK NDMVEQMHED IISLWDQSLK PCVKLTPLCV SLKCTDL-CN ATNTNSSNTN SSSGEMMME-
ARV 2             C            N     Q                        T N     - K        --- ----NWKE I
LAV MAL         IE E      G    N                              T N  NVN T  V GTNACS  RTNA LK  I
LAV ELI         IA E           N                              T N S E--L RN GTMG NV TTEEKG----

                170        180        190        200        210        220        230        240
LAV BRU  KGEIKNCSFN ISTSIRGKVQ KEYAFFYKLD IIPIDNDTTS -----YTLTS CNTSVITQAC PKVSFEPIPI HYCAPAGFAI
ARV 2                T    D I    N L RN  VV     AS T TNYTN R IN    R                     T
LAV MAL  -  V         TPVGSD R  -    T  N  LVQ  DSDN  ----S R IN                  T D
LAV ELI  ---M        VT VLKD K  QV   L  R     V     SST -NSTN R IN      A

                250        260        270        280        290        300        310        320
LAV BRU  LKCNKKTFNG TGPCTNVSTV QCTHGIRPVV STQLLLNGSL AEEEVVIRSA NFTDNAKTII VQLNQSVEIN CTRPNNNTRK
ARV 2                K                 I                        D   N          E  A
LAV MAL     D K       EI K          K                   IM      E   L   T N       ET T      G    R
LAV ELI    RD K                                          I      E   L N   N   AH  E   K T  A  YQ   Q
```

FIG. 3E-1

```
                330        340        350        360        370        380        390        400
LAV BRU  SIRIQRGPGR AFVTIGK-IG NMRQAHCNIS RAKWNATLKQ IASKLREQFG NNKT-IIFKQ SSGGDPEIVT HSFNCGGEFF
ARV 2      Y --        W T RI   DI K         Q  N  E   VK             - V N          M     R
LAV MAL  G HF--    Q  LY T I-V  DI R Y T N ETE DK  Q   V V  GSLL- -   K   NS         T     R
LAV ELI  RTP -- L Q SLY TKS-RS IIG           Q SK  Q   V R  GTLL- -   I K  P         T

                410        420        430        440        450        460        470        480
LAV BRU  YCNSTQLFNS TWFNSTWSTE CSNNTEGSDT ITLPCRIKQF INMWQEVGKA MYAPPISGQI RCSSNITGLL LTRDGGNN--
ARV 2      T     N    -----RLN RTEG K N    I       I                      C   S               T -V
LAV MAL    TSK        Q NGARL- -  S  STGS          I      KT             A V  N L      I      NSSD
LAV ELI    TSG        NI A NNI TES NSTNTN    Q     I  K VAGR-    I       ERN  L                I --

                490        500        510        520        530        540        550        560
LAV BRU  NNGSEIFRPG GGDMRDNWRS ELYKYKVVKI EPLGVAPTKA KRRVVQREKR AVGI-GALFL GFLGAAGSTM GARSMTLTVQ
ARV 2    T DT V                      I       I                       V  M                 V L
LAV MAL   SDN TL              I      R                     E      I L-  M                 A L
LAV ELI   STN T                      Q          R          E      I L-  M                   V
```

FIG. 3E-2

```
                570        580        590        600        610        620        630        640
LAV BRU  ARQLLSGIVQ QQNNLLRAIE AQQHLLQLTV WGIKQLQARI LAVERYLKDQ QLLGIWGCSG KLICTTAVPW NASWSNKSLE
ARV 2                                              W        R
LAV MAL                                            W        Q   R   M       H    F     S    R  D
LAV ELI      M                                                              H    N     S    R  N

                650        660        670        680        690        700        710        720
LAV BRU  QIWNNMTWME WDREINNYTS LINSLIEESQ NQQEKNEQEL LELDKWASLW NWFNITNWLW YIKIFIMIVG GLVGLRIVFA
ARV 2    D  D     Q  E   D   N T YT L                              S
LAV MAL  D        Q  EK  S   G I YN       I      K                 S SK      R   IV     I    I
LAV ELI  E  Q        E   D   G   Y        T      K                 S  Q             I   I

                730        740        750        760        770        780        790        800
LAV BRU  VLSIVNRVRQ GYSPLSFQTH LPTPRGP-DR PEGIEEEGGE RDRDRSIRLV NGSLALIWDD LRSLCLFSYH RLRDLLLIVT
ARV 2                        R   V    -    D               V    D F     E           R         AA
LAV MAL     L             L  L        P              QG G         FS         N                 A
LAV ELI     L                L   A    -      T        G    V  L   FS                       I  AV

                810        820        830        840        850        860        870
LAV BRU  RIVELLGRRG WEALKYWWNL LQYWSQELKN SAVSLLNATA IAVAEGTDRV IEVVQGACRA IRHIPRRIRQ GLERILL
ARV 2     T  I  K           S      I          W         T          A R Y    L  H          L
LAV MAL                   L        G        I    T        C       IG RFG    L          F -A
LAV ELI               DI  L            R    S   FD I              II R     VLN            S
```

FIG. 3F-1

F

```
                   10         20         30         40         50         60         70         80
LAV BRU  MGGKWSKSSV VGWPTVRERM R----RAEPA ADGVGAASR- ----DLEKUG AITSSNTAAT NAACAWLEAQ EE-EEVGFPV
ARV 2          R M G  SAI     RAEP              V  - ----                       D          -
LAV MAL            I     KI    I  ---- TP T ET   V QD AVSQ  D C   AA  SP  N   S --- PP   E
LAV ELI            I     AI    I  ---- TM        V  - ----                    S   D          SD

                   90        100        110        120        130        140        150        160
LAV BRU  TPQVPLRRHT YKAAVDLSHF LKEKGGLEGL IHSQRRQDIL DLWIYUTQGY FPDWQNYTPC PGVRYPLTFG WCYKLVPVEP
ARV 2    R              L  I             W     E                           I          F
LAV MAL  R             G F          D   VW PK  E      V                     I F        F    HS
LAV ELI  R             E L                W KK  E      V N    I              I           E    D

                  170        180        190        200        210
LAV BRU  DKVEEANKGE NTSLLHPVSL HGMDDPEREV LEWRFDSRLA FHHVARELHP EYFKNC
ARV 2    E      E    N     M      E A K   V    K      M          Y D
LAV MAL  EE     E    NC    I Q    E A     K K  S   LR R  Q       Y D
LAV ELI  QE   DTE    TN    ICQ    E    Q  K   N      E K  M     FY  -
```

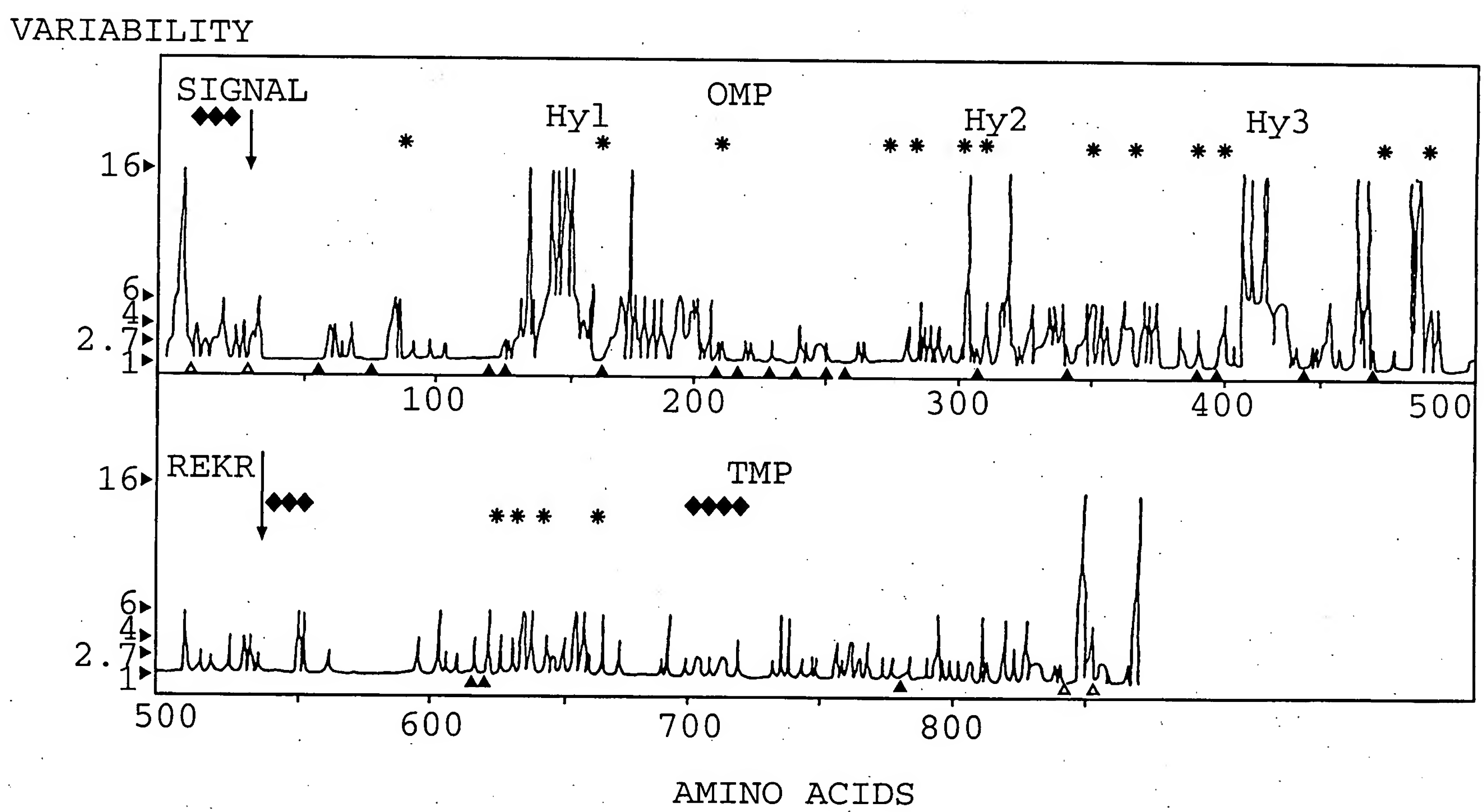
FIG. 3F-2

A LAVbru vs.	GAG		POL		ENV					
					TOTAL		OMP		TMP	
HTLV-3 USA	512 0/0	0.8	1015 0/0	1.3	856 5/0	1.4	507 5/0	1.6	349 0/0	1.1
ARV-2 USA	502 12/2	3.4	1003 12/0	3.1	855 17/11	13.0	505 17/10	14.3	350 0/1	11.2
LAVeli ZAIRE	500 13/1	9.8	1002 13/0	5.5	853 22/14	20.7	504 22/14	25.3	349 0/0	13.8
LAVmal ZAIRE	505 14/7	12.0	1002 13/0	7.7	859 13/11	21.7	509 13/10	26.4	350 0/1	14.9
B LAVeli vs.										
LAVmal	505 1/6	10.8	1002 0/0	8.4	859 13/11	19.8	509 8/13	23.6	350 0/1	14.3

FIG. 4A

A LAVbru vs.	orf F		central region					
			orf Q		orf R		orf S	
HTLV-3 USA	206 0/0	1.5	192 0/0	0		nd	80 0/0	2.5
ARV-2 USA	210 0/4	12.6	192 0/0	10.0	97 0/1	9.4	81 0/1	15.0
LAVeli ZAIRE	206 1/1	19.4	192 0/0	10.4	96 0/0	11.5	80 0/0	27.5
LAVmal ZAIRE	209 2/5	27.0	192 0/0	12.6	96 0/0	10.4	80 0/0	23.8
B LAVeli vs.								
LAVmal	209 3/6	22.5	192 0/0	12.0	96 0/0	6.3	80 0/0	11.3

FIG. 4B



VARIABILITY
SIGNAL
Hy1
OMP
Hy2
Hy3
REKR
TMP
16
6
4
2.7
1
100
200
300
400
500
600
700
800
AMINO ACIDS

FIG. 5

GAG

a

```
                                 120

          K   A   Q   Q   A   A   A                           D   T
LAV.BRU  AAA GCA CAG CAA GCA GCA GCT  -   -   -   -   -      GAC ACA

          K   A   Q   Q   A   A   A   A   A   -   -   -   -   G   T
ARV 2    AAG GCA CAG CAA GCA GCA GCT GCA GCT                 GGC ACA

          K   T   Q   Q   A   A   A   A   Q   Q   A   A   A   A   T
LAV.MAL  AAG ACA CAG CAG GCA GCA GCT GCA GAG GAG GCA GCA GCT GCC ACA

          X   A   Q   Q   A   A   A                           D   T
LAV.ELI  AAG GCA CAG CAA GCA GCA GCT  -   -   -   -   -   -  GAC ACA
```

FIG. 6A-1

b

LAV.BRU

```
                                460                                                 470                                                 480
G   N   F   L   Q   S   R   P   E   P   T   A   P   P   F   L   Q   S   R   P   E   P   T   A   P   P   E   E
GGG AAT TTT CTT CAG AGC AGA CCA GAG CCA ACA GCC CCA CCA TTT CTT CAG AGC AGA CCA GAG CCA ACA GCC CCA CCA GAA GAG
```

ARV 2

```
G   N   F   L   Q   S   R   P   E   P   T   A   P   P                                                   E   E
GGG AAT TTT CTT CAG AGC AGA CCA GAG CCA ACA GCC CCA CCA - - - - - - - - - - - - - - -                   GAA GAG
```

LAV.MAL

```
G   N   F   L   Q   S   R   P   E   P   T   A   P   P                                                   A   E
GGG AAT TTC CTT CAG AGC AGA CCA GAG CCA ACA GCC CCA CCA - - - - - - - - - - - - - - -                   GCA GAG
```

LAV.ELI

```
G   N   F   L   Q   S   R   P   E   P   T   A   P   P                                                   A   E
GGG AAC TTT CTC CAA AGC AGA CCA GAG CCA ACA GCC CCA CCA - - - - - - - - - - - - - - -                   GCA GAG
```

FIG. 6A-2

```
c
                          20                                      30
               R    M    R                        R    A    E    P    A
LAV.BRU       AGA  ATG  AGA   -    -    -    -   CGA  GCT  GAG  CCA  GCA

               R    M    R    R    A    E    P    R    A    E    P    A
ARV 2         AGA  ATG  AGA  CGA  GCT  GAG  CCA  CGA  GCT  GAG  CCA  GCA

               R    I    R                        R    T    P    P    T
LAV.MAL       AGA  ATA  AGA   -    -    -    -   CGA  ACT  CCC  CCA  ACA

               R    I    R    -    -    -    -    R    T    P    P    T
LAV.ELI       AGA  ATA  AGA                      AGA  ACT  AAT  CCA  GCA

d
                                         40
               V    G    A    A    S    R                             D
LAV.BRU       GTG  GGA  GCA  GCA  TCT  CGA   -    -    -    -    -   GAC

               V    G    A    V    A    R                             D
ARV 2         GTG  GGA  GCA  GTA  TCT  CGA   -    -    -    -    -   GAC

               V    G    A    V    S    R    D    A    V    S    Q    D
LAV.MAL       GTA  GGA  GCA  GTA  TCT  CAA  GAT  GCA  GTA  TCT  CAA  GAT

               V    G    A    V    S    R                             D
LAV.ELI       GTA  GGA  GCA  GTA  TCT  CGA   -    -    -    -    -   GAC
```

FIG. 6A-3

```
e                                      ENV
                                          20
           Q   H   L   W   R   W   G   W   K   W   G   T   M   L
LAV.BRU   CAG CAC TTG TGG ACA TGG GGC TGG AAA TGG GGC ACC ATG CTC

           Q   H   L   W   R   W   G                   T   L   L
ARV 2     CAG CAC TTG TGG AGA TGG GGC  -   -   -   -  ACC TTG CTC
           Q   N   W   W   R   W   G                   M   M   L
LAV.MAL   CAA AAC TGG TGG AGA TGG GGC  -   -   -   -  ATG ATG CTC
           Q   N   W   W   K   W   G                   T   M   L
LAV.ELI   CAA AAC TGG TGG AAA TCG GGC  -   -   -   -  ATC ATG CTC

f
LAV.BRU                          140                                       150
 L   K   C   T   D   L       G   N   A   T   N   T   N   S   S   N   T   N   S   S   S   G   E
TTA AAG TGC ACT GAT TTG  -  GGG AAT GCT ACT AAT ACC AAT AGT AGT AAT ACC AAT AGT AGT AGC GGG GAA
 M   M   M   E       K   G   E   I
ATG ATG ATG GAG  -  AAA GCA GAG ATA
ARG 2
 L   N   C   T   D   L       G   K   A   T   N   T   N   S   S                               M
TTA AAT TGC ACT GAT TTG  -  GGG AAG GCT ACT AAT ACC AAT AGT AGT  -   -   -   -   -   -   -  AAT
 W   K   E   E   I   K   G   E   I
TGG AAA GAA GAA ATA AAA GGA GAA ATA
```

FIG. 6B-1

LAV.MAL

```
L   N   C   T   N   V   N   G   T   A   V   N   G   T   N   A   G   S   N   R   T   N   A   E
TTA AAC TGC ACT AAT GTG AAT GGG ACT GCT GTG AAT GGG ACT AAT GCT GGG AGT AAT AGG ACT AAT GCA GAA

L   K   M   E   I       G   E   V
TTG AAA ATG GAA ATT -   GGA GAA GTG
```

LAV.ELI

```
L   N   C   S   D   E           L   R   N   N   G   T   M   G   N   N   V   T   T   E   E   K
TTA AAC TGT AGT GAT GAA -   -   TTG AGG AAC AAT GGC ACT ATG GGG AAC AAT GTC ACT ACA GAG GAG AAA

G                               M
GGA -   -   -   -   -   -   -   ATG
```

FIG. 6B-2

```
g                                      200
           D   N   D   T   T   S                          Y   T   L
LAV.BRU   GAT AAT GAT ACT ACC AGC  -   -   -   -   -     TAT ACG TTG

           D   N   A   S   T   T   T   N   Y   T   N   Y   R   L
ARV 2     GAT AAT GCT AGT ACT ACT ACC AAC TAT ACC AAC TAT AGG TTG

           D   D   S   D   N   S   S                      Y   R   L
LAV.MAL   GAT GAT AGT GAT AAT AGT AGT  -   -   -   -     TAT AGG CTA

           D   N   D   S   S   T       N   S   T   N   Y   R   L
LAV.ELI   GAC AAT GAT AGT AGT ACC  -  AAT AGT ACC AAT TAT AGG TTA

h

LAV.BRU                      410                                            420                 430

 C   N   S   T   Q   L   F   N   S   T   W   F   N   S   T   W   S   T   E   G   S   N   N   T   E   G
TGT AAT TCA ACA CAA CTG TTT AAT AGT ACT TGG TTT AAT AGT ACT TGG AGT ACT GAA GGG TCA AAT AAC ACT GAA GGA

 S   D   T   I
AGT GAC ACA ATC

ARV 2

 C   N   T   T   Q   L   F   N   N   T   W                       R   L   N   H   T   K   G   T   K   G
TGT AAT ACA ACA CAA CTG TTT AAT AAT ACA TGG  -   -   -   -   -  AGG TTA AAT CAC ACT GAA GGA ACT AAA GGA

 N   D   T   I
AAT GAC ACA ATC
```

FIG. 6B-3

LAV.MAL

```
 C   N   T   S   K   L   F   N   S   T   W   Q   N   N   G   A   R   L           S   N   S   T   E   S
TGT AAT ACA TCA AAA CTG TTT AAT AGT ACA TGG CAG AAT AAT GGT GCA AGA CTA  -   -  AGT AAT AGC ACA GAG TCA

 T   G   S   I
ACT GGT AGT ATC
```

LAV.ELI

```
 C   N   T   S   G   L   F   N   S   T   W   N   I   S   A   W   N   N   I   T   E   S   N   N   S   T
TGT AAT ACA TCA GGA CTG TTT AAT AGT ACA TGG AAT ATT AGT GCA TGG AAT AAT ATT ACA GAG TCA AAT AAT AGC ACA

 N   T   N   I
AAC ACA AAC ATC
```

FIG. 6B-4

LAV.MAL

R
GGTCTCTCTTGTTAGACCAGGTCGAGCCCGGGAGCTCTCTGGCTAGCAAGGAACCCACTG

R U5
CTTAAGCCTCAATAAAGCTTGCCTTGAGTGCCTCAAGCAGTGTGTGCCCATCTGTTGTGT

100 U5
GACTCTGGTAACTAGAGATCCCTCAGACCACTCTAGACGGTGTAAAAATCTCTAGCAGTG

GCGCCCGAACAGGGACTTTAAAGTGAAAGTAACAGGGACTCGAAAGCGGAAGTTCCAGAG

200
AAGTTCTCTCGACGCAGGACTCGGCTTGCTGAGGTGCACACAGCAAGAGGCGAGAGCGGC

GAG 300
MetGlyAlaArg
GACTGGTGAGTACGCCAATTTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAG

AlaSerValLeuSerGlyGlyLysLeuAspAlaTrpGluLysIleArgLeuArgProGly
AGCGTCAGTATTAAGCGGGGGAAAATTAGATGCATGGGAGAAAATTCGGTTAAGGCCAGG

400
GlyLysLysLysTyrArgLeuLysHisLeuValTrpAlaSerArgGluLeuGluArgPhe
GGGAAAGAAAAAATATAGACTGAAACATTTAGTATGGGCAAGCAGGGAGCTGGAAAGATT

AlaLeuAsnProGlyLeuLeuGluThrGlyGluGlyCysGlnGlnIleMetGluGlnLeu
CGCACTTAACCCTGGCCTTTTAGAAACAGGAGAAGGATGTCAACAAATAATGGAACAGCT

500
GlnSerThrLeuLysThrGlySerGluGluIleLysSerLeuTyrAsnThrValAlaThr
ACAATCAACTCTCAAGACAGGATCAGAAGAAATTAAATCATTATATAATACAGTAGCAAC

600
LeuTyrCysValHisGlnArgIleAspValLysAspThrLysGluAlaLeuAspLysIle
CCTCTATTGTGTACATCAAAGGATAGATGTAAAAGACACCAAGGAAGCGCTAGATAAAAT

GluGluIleGlnAsnLysSerArgGlnLysThrGlnGlnAlaAlaAlaAlaGlnGlnAla
AGAGGAAATACAAAATAAGAGCAGGCAAAAGACACAGCAGGCAGCAGCTGCACAGCAGGC

700
AlaAlaAlaThrLysAsnSerSerSerValSerGlnAsnTyrProIleValGlnAsnAla
AGCAGCTGCCACAAAAAACAGCAGCAGTGTCAGTCAAAATTACCCCATAGTGCAAAATGC

GlnGlyGlnMetIleHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysVal
ACAAGGGCAAATGATACATCAGGCCATATCACCTAGGACTTTGAATGCATGGGTGAAAGT

800
IleGluGluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGly
AATAGAAGAAAAGGCTTTCAGCCCAGAAGTGATACCCATGTTCTCAGCATTATCAGAGGG

900
AlaThrProGlnAspLeuAsnMetMetLeuAsnIleValGlyGlyHisGlnAlaAlaMet
GGCCACCCCACAAGATTTAAATATGATGCTGAACATAGTTGGAGGACACCAGGCAGCTAT

GlnMetLeuLysAspThrIleAsnGluGluAlaAlaAspTrpAspArgValHisProVal
GCAAATGTTAAAAGATACCATCAATGAGGAAGCTGCAGACTGGGACAGGGTACATCCAGT

1000
HisAlaGlyProIleProProGlyGlnMetArgGluProArgGlySerAspIleAlaGly
ACATGCAGGGCCTATTCCCCCAGGCCAGATGAGAGAACCAAGAGGAAGTGACATAGCAGG

FIG. 7A

```
 ThrThrSerThrLeuGlnGluGlnIleGlyTrpMetThrSerAsnProProIleProVal
AACTACTAGTACCCTTCAAGAACAAATAGGATGGATGACAAGCAACCCACCTATCCCAGT
              1100
 GlyAspIleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSer
GGGAGACATCTATAAAAGATGGATAATCCTGGGATTAAATAAAATAGTAAGAATGTATAG
                                                         1200
 ProValSerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAsp
CCCTGTCAGCATTTTGGACATAAGACAAGGGCCAAAGGAACCTTTTAGAGACTATGTAGA

 ArgPhePheLysThrLeuArgAlaGluGlnAlaThrGlnGluValLysAsnTrpMetThr
TAGGTTCTTTAAAACTCTCAGAGCTGAGCAAGCTACACAGGAGGTAAAAAATTGGATGAC
                                            1300
 GluThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGly
AGAAACCTTGCTGGTCCAAAATGCGAATCCAGACTGTAAGACCATTTTAAAAGCATTAGG

 ProGlyAlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProSerHis
ACCAGGGGCTACATTAGAAGAAATGATGACAGCATGCCAGGGAGTGGGAGGACCCAGTCA
              1400
 LysAlaArgValLeuAlaGluAlaMetSerGlnAlaThrAsnSerThrAlaAlaIleMet
TAAAGCAAGAGTTTTGGCTGAGGCAATGAGCCAAGCAACAAATTCAACTGCTGCCATAAT
                                                         1500
 MetGlnArgGlyAsnPheLysGlyGlnLysArgIleLysCysPheAsnCysGlyLysGlu
GATGCAGAGAGGTAATTTTAAGGGCCAGAAAAGAATTAAGTGTTTCAACTGTGGCAAAGA

 GlyHisLeuAlaArgAsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLys
AGGACACCTAGCCAGAAATTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAAATGTGGAA
                                            1600  →POL
                                                  PhePheArgGluAsnLeu
 GluGlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrp
GGAAGGACACCAAATGAAAGACTGCACTGAGAGACAGGCTAATTTTTTAGGGAAAATTTG

AlaPheProGlnGlyLysAlaArgGluPheProSerGluGlnThrArgAlaAsnSerPro
 ProSerHisLysGlyArgProGlyAsnPheLeuGlnSerArgProGluProThrAlaPro
GCCTTCCCACAAGGGAAGGCCAGGGAATTTCCTTCAGAGCAGACCAGAGCCAACAGCCCC
              1700
ThrSerArgGluLeuArgValTrpGlyGlyAspLysThrLeuSerGluThrGlyAlaGlu
 ProAlaGluSerPheGlyPheGlyGluGluIleLysProSerGlnLysGlnGluGlnLys
ACCAGCAGAGAGCTTCGGGTTTGGGGAGGAGATAAAACCCTCTCAGAAACAGGAGCAGAA
                                                         1800
ArgGlnGlyIleValSerPheSerPheProGlnIleThrLeuTrpGlnArgProValVal
 AspLysGluLeuTyrProLeuAlaSerLeuLysSerLeuPheGlyAsnAspGlnLeuSer
AGACAAGGAATTGTATCCTTTAGCTTCCCTCAAATCACTCTTTGGCAACGACCAGTTGTC
GAG ←
ThrValArgValGlyGlyGlnLeuLysGluAlaLeuLeuAspThrGlyAlaAspAspThr
 Gln
ACAGTAAGAGTAGGAGGACAGCTAAAAGAAGCTCTATTAGACACAGGAGCAGATGATACA
                                            1900
ValLeuGluGluIleAsnLeuProGlyLysTrpLysProLysMetIleGlyGlyIleGly
GTATTAGAAGAAATAAATTTGCCAGGAAAATGGAAACCAAAAATGATAGGGGGAATTGGA

GlyPheIleLysValArgGlnTyrAspGlnIleLeuIleGluIleCysGlyLysLysAla
GGTTTTATCAAAGTAAGACAGTATGATCAAATACTTATAGAAATTTGTGGAAAAAAGGCT
              2000
```

FIG. 7B

IleGlyThrIleLeuValGlyProThrProValAsnIleIleGlyArgAsnMetLeuThr
ATAGGTACAATATTGGTAGGACCTACACCTGTCAACATAATTGGACGAAATATGTTGACT
2100
GlnIleGlyCysThrLeuAsnPheProIleSerProIleGluThrValProValLysLeu
CAGATTGGTTGTACTTTAAATTTTCCAATTAGTCCTATTGAGACTGTACCAGTAAAATTA

LysProGlyMetAspGlyProArgValLysGlnTrpProLeuThrGluGluLysIleLys
AAGCCAGGGATGGATGGCCCAAGGGTTAAACAATGGCCATTGACAGAAGAAAAAATAAAA
2200
AlaLeuThrGluIleCysLysAspMetGluLysGluGlyLysIleLeuLysIleGlyPro
GCATTAACAGAAATTTGTAAAGATATGGAAAAGGAAGGAAAAATTTTAAAAATTGGGCCT

GluAsnProTyrAsnThrProValPheAlaIleLysLysLysAspSerThrLysTrpArg
GAAAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGCACTAAATGGAGA
2300
LysLeuValAsnPheArgGluLeuAsnLysArgThrGlnAspPheTrpGluValGlnLeu
AAATTAGTGAATTTCAGAGAGCTTAATAAAAGAACTCAAGATTTTTGGGAAGTTCAATTA
2400
GlyIleProHisProAlaGlyLeuLysLysLysLysSerValThrValLeuAspValGly
GGAATACCACATCCTGCTGGGTTGAAAAAGAAAAAATCAGTCACAGTATTGGATGTGGGG

AspAlaTyrPheSerValProLeuAspGluAspPheArgLysTyrThrAlaPheThrIle
GATGCATATTTTTCAGTCCCTTTAGATGAAGATTTCAGGAAGTATACTGCATTCACTATA
2500
ProSerIleAsnAsnGluThrProGlyIleArgTyrGlnTyrAsnValLeuProGlnGly
CCCAGTATTAATAATGAGACACCAGGGATTAGATATCAGTACAATGTGCTACCACAGGGA

TrpLysGlySerProAlaIlePheGlnSerSerMetThrLysIleLeuGluProPheArg
TGGAAAGGATCACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAACCCTTT AGA
2600
ThrLysAsnProGluIleValIleTyrGlnTyrMetAspAspLeuTyrValGlySerAsp
ACAAAAAATCCAGAAATAGTCATATACCAATACATGGATGATTTGTATGTAGGGTCTGAT
2700
LeuGluIleGlyGlnHisArgThrLysIleGluGluLeuArgGluHisLeuLeuLysTrp
TTAGAAATAGGACAACATAGAACAAAAATAGAGGAACTAAGAGAACATCTATTGAAATGG

GlyPheThrThrProAspLysLysHisGlnLysGluProProPheLeuTrpMetGlyTyr
GGATTTACCACACCAGACAAAAAGCATCAGAAAGAACCCCCATTTCTTTGGATGGGGTAT
2800
GluLeuHisProAspLysTrpThrValGlnProIleGlnLeuProAspLysGluSerTrp
GAACTCCACCCTGACAAATGGACAGTGCAGCCTATACAACTGCCAGACAAGGAAAGCTGG

ThrValAsnAspIleGlnLysLeuValGlyLysLeuAsnTrpAlaSerGlnIleTyrPro
ACTGTCAATGATATACAGAAATTGGTGGGAAAACTAAATTGGGCAAGTCAGATTTATCCA
2900
GlyIleLysValLysGlnLeuCysLysLeuLeuArgGlyAlaLysAlaLeuThrAspIle
GGAATTAAAGTAAAGCAATTATGTAAACTCCTTAGGGGAGCAAAAGCACTAACAGACATA
3000
ValProLeuThrAlaGluAlaGluLeuGluLeuAlaGluAsnArgGluIleLeuLysGlu
GTACCATTAACTGCAGAGGCAGAATTAGAATTGGCAGAGAACAGGGAAATTCTAAAAGAA

FIG. 7C

ProValHisGlyValTyrTyrAspProSerLysAspLeuIleAlaGluIleGlnLysGln
CCAGTGCATGGGGTATATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGAAGCAG
3100
GlyGlnGlyGlnTrpThrTyrGlnIleTyrGlnGluGlnTyrLysAsnLeuLysThrGly
GGGCAAGGTCAATGGACATATCAAATATACCAAGAGCAATATAAAAATCTGAAAACAGGG

LysTyrAlaArgIleLysSerAlaHisThrAsnAspValLysGlnLeuThrGluAlaVal
AAGTATGCAAGAATAAAGTCTGCCCACACTAATGATGTAAAACAATTAACAGAAGCAGTG
3200
GlnLysIleAlaGlnGluSerIleValIleTrpGlyLysThrProLysPheArgLeuPro
CAAAAGATAGCCCAAGAAAGCATAGTAATATGGGGAAAAACTCCTAAATTTAGACTACCC
3300
IleGlnLysGluThrTrpGluAlaTrpTrpThrGluTyrTrpGlnAlaThrTrpIlePro
ATACAAAAAGAAACATGGGAGGCATGGTGGACAGAATATTGGCAAGCCACCTGGATCCCT

GluTrpGluPheValAsnThrProProLeuValLysLeuTrpTyrGlnLeuGluThrGlu
GAATGGGAGTTTGTCAATACTCCTCCCCTAGTAAAACTATGGTACCAGTTAGAAACAGAA
3400
ProIleValGlyAlaGluThrPheTyrValAspGlyAlaAlaAsnArgGluThrLysLys
CCCATAGTAGGAGCAGAAACTTTCTATGTAGATGGGGCAGCTAATAGAGAAACTAAAAAG

GlyLysAlaGlyTyrValThrAspArgGlyArgGlnLysValValSerLeuThrGluThr
GGAAAAGCAGGATATGTTACTGACAGAGGAAGACAAAAGGTTGTCTCCTTAACTGAAACA
3500
ThrAsnGlnLysThrGluLeuGlnAlaIleHisLeuAlaLeuGlnAspSerGlySerGlu
ACAAATCAGAAGACTGAATTACAAGCAATCCACTTAGCTTTACAGGATTCAGGATCAGAA
3600
ValAsnIleValThrAspSerGlnTyrAlaLeuGlyIleIleGlnAlaGlnProAspLys
GTAAACATAGTAACAGACTCACAGTATGCATTAGGGATTATTCAAGCACAACCAGATAAA

SerGluSerGluIleValAsnGlnIleIleGluGlnLeuIleGlnLysAspLysValTyr
AGTGAATCAGAGATTGTTAATCAAATAATAGAGCAATTAATACAGAAGGACAAGGTCTAC
3700
LeuSerTrpValProAlaHisLysGlyIleGlyGlyAsnGluGlnValAspLysLeuVal
CTGTCATGGGTACCAGCACACAAAGGGATTGGAGGAAATGAACAAGTAGATAAATTAGTC

SerSerGlyIleArgLysValLeuPheLeuAspGlyIleAspLysAlaGlnGluGluHis
AGCAGTGGAATCAGAAAGGTACTATTTTTAGATGGGATAGATAAGGCTCAAGAAGAACAT
3800
GluLysTyrHisSerAsnTrpArgAlaMetAlaSerAspPheAsnLeuProProIleVal
GAAAAATATCACAGCAATTGGAGAGCAATGGCTAGTGACTTTAATCTACCACCTATAGTA
3900
AlaLysGluIleValAlaSerCysAspLysCysGlnLeuLysGlyGluAlaMetHisGly
GCGAAGGAAATAGTAGCCAGCTGTGATAAATGTCAACTAAAAGGGGAAGCCATGCATGGA

GlnValAspCysSerProGlyIleTrpGlnLeuAspCysThrHisLeuGluGlyLysIle
CAAGTAGACTGTAGTCCAGGGATATGGCAATTAGATTGCACACATCTAGAAGGAAAAATA
4000
IleIleValAlaValHisValAlaSerGlyTyrIleGluAlaGluValIleProAlaGlu
ATCATAGTAGCAGTCCATGTAGCCAGTGGATATATAGAAGCAGAAGTTATCCCAGCAGAA

ThrGlyGlnGluThrAlaTyrPheIleLeuLysLeuAlaGlyArgTrpProValLysVal
ACAGGACAGGAGACAGCATACTTTATACTAAAATTAGCAGGAAGATGGCCAGTAAAAGTA
4100

FIG. 7D

```
ValHisThrAspAsnGlySerAsnPheThrSerAlaAlaValLysAlaAlaCysTrpTrp
GTACACACAGACAATGGCAGCAATTTCACCAGTGCTGCAGTTAAAGCAGCCTGTTGGTGG
                                                        4200
AlaAsnIleLysGlnGluPheGlyIleProTyrAsnProGlnSerGlnGlyValValGlu
GCAAATATCAAACAGGAATTTGGAATTCCCTACAACCCCCAAAGTCAAGGAGTAGTGGAA

SerMetAsnLysGluLeuLysLysIleIleGlyGlnValArgGluGlnAlaGluHisLeu
TCTATGAATAAGGAATTAAAGAAAATCATAGGGCAGGTAAGAGAGCAAGCTGAACACCTT
                                    4300
LysThrAlaValGlnMetAlaValPheIleHisAsnPheLysArgLysGlyGlyIleGly
AAGACAGCAGTACAAATGGCAGTGTTCATTCACAATTTTAAAAGAAAAGGGGGGATTGGG

GlyTyrSerAlaGlyGluArgIleIleAspMetIleAlaThrAspIleGlnThrLysGlu
GGGTACAGTGCAGGGGAAAGAATAATAGACATGATAGCAACAGACATACAAACTAAAGAA
                  4400
LeuGlnLysGlnIleThrLysIleGlnAsnPheArgValTyrTyrArgAspAsnArgAsp
TTACAAAAACAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAGGGACAACAGAGAC
                                                        4500
ProIleTrpLysGlyProAlaLysLeuLeuTrpLysGlyGluGlyAlaValValIleGln
CCAATTTGGAAAGGACCAGCAAAACTACTCTGGAAAGGTGAAGGGGCAGTAGTAATACAG
                                                          →Q
AspAsnSerAspIleLysValValProArgArgLysAlaLysIleIleArgAspTyrGly
                                                         MetGlu
GACAATAGTGATATAAAGGTAGTACCAAGAAGAAAAGCAAAAATCATTAGGGATTATGGA
                                    4600  POL←
LysGlnMetAlaGlyAspAspCysValAlaGlyGlyGlnAspGluAsp
 AsnArgTrpGlnValMetIleValTrpGlnValAspArgMetArgIleArgThrTrpHis
AAACAGATGGCAGGTGATGATTGTGTGGCAGGTGGACAGGATGAGGATTAGAACATGGCA

 SerLeuValLysHisHisMetTyrValSerLysLysAlaLysAsnTrpPheTyrArgHis
CAGTTTAGTAAAACATCATATGTATGTCTCAAAGAAAGCTAAAAATTGGTTTTATAGACA
          4700
 HisTyrGluSerArgHisProLysValSerSerGluValHisIleProLeuGlyAspAla
TCACTATGAAAGCAGGCATCCAAAAGTAAGTTCAGAAGTACACATCCCACTAGGGGATGC
                                                        4800
 ArgLeuValValArgThrTyrTrpGlyLeuGlnThrGlyGluLysAspTrpHisLeuGly
TAGATTAGTAGTAAGAACATATTGGGGTCTGCAAACAGGAGAAAAAGACTGGCACTTGGG

 HisGlyValSerIleGluTrpArgGlnLysArgTyrSerThrGlnLeuAspProAspLeu
TCATGGGGTCTCCATAGAATGGAGGCAGAAAAGATATAGCACACAACTAGATCCTGACCT
                                    4900
 AlaAspGlnLeuIleHisLeuTyrTyrPheAspCysPheSerGluSerAlaIleArgGln
AGCAGACCAACTGATTCATCTGTACTATTTTGATTGTTTTTCAGAATCTGCCATAAGACA

 AlaIleLeuGlyHisIleValSerProArgCysAspTyrGlnAlaGlyHisAsnLysVal
AGCCATATTAGGACATATAGTTAGTCCTAGGTGTGATTATCAAGCAGGACATAACAAGGT
                  5000
 GlySerLeuGlnTyrLeuAlaLeuThrAlaLeuIleAlaProLysLysThrArgProPro
AGGATCTTTACAGTATTTGGCACTAACAGCATTAATAGCACCAAAAAAGACAAGGCCACC
                                   →R                   5100
                                   MetGluGlnAlaProAlaAspGlnGly
 LeuProSerValArgLysLeuThrGluAspArgTrpAsnLysProGlnGlnThrLysGly
TTTGCCTAGTGTTAGGAAGCTAACAGAAGATAGATGGAACAAGCCCCAGCAGACCAAGGG
```

FIG. 7E

```
                              Q←─┐
ProGlnArgGluProHisAsnGluTrpThrLeuGluLeuLeuGluGluLeuLysGlnGlu
 HisArgGlySerHisThrMetAsnGlyHis
CCACAGAGGGAGCCACACAATGAATGGACATTAGAACTTTTAGAGGAGCTTAAGCAAGAA
                                        5200
AlaValArgHisPheProArgIleTrpLeuHisSerLeuGlyGlnHisIleTyrGluThr
GCTGTCAGACACTTTCCTAGGATATGGCTCCATAGTTTAGGACAACATATCTATGAAACT

TyrGlyAspThrTrpGluGlyValGluAlaIleIleArgSerLeuGlnGlnLeuLeuPhe
TATGGGGATACCTGGGAAGGAGTTGAAGCTATAATAAGAAGTCTGCAACAACTGCTGTTT
            5300
IleHisPheArgIleGlyCysGlnHisSerArgIleGlyIleThrArgGlnArgArgAla
ATTCATTTCAGAATTGGGTGTCAACATAGCAGAATAGGCATTACTCGACAGAGAAGAGCA
   ┌→S      R←─┐                                         5400
ArgAsnGlySerSerArgSer
   MetAspProValAspProAsnLeuGluProTrpAsnHisProGlySerGlnProArg
AGAAATGGATCCAGTAGATCCTAACTTAGAGCCCTGGAACCATCCAGGGAGTCAGCCTAG

 ThrProCysAsnLysCysTyrCysLysLysCysCysTyrHisCysGlnMetCysPheIle
GACGCCTTGTAATAAGTGTTATTGTAAAAAGTGCTGCTATCATTGCCAAATGTGCTTCAT
                                        5500
 ThrLysGlyLeuGlyIleSerTyrGlyArgLysLysArgArgGlnArgArgArgProPro
AACGAAAGGCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGACCTCC
                                      S←─┐
 GlnGlyAsnGlnAlaHisGlnAspProLeuProGluGln
TCAGGGCAATCAGGCTCATCAAGATCCTCTACCAGAGCAGTAAGTAGTATATGTAATACA
            5600
ACCTTTAGTGATATTAGCAATAGTAGCATTAGTAGTAACGCTAATAATAGCAATAGTTGT
                                                        5700
GTGGACCATAGTATTTATAGAAATTAGGAAAATAAGAAGACAAAGGAAAATAGACAGGTT
                                      ┌→ENV
                                      MetArgValArgGluIleGlnArg
GATTGATAGAATAAGAGAAAGAGCAGAAGATAGTGGCAATGAGAGTGAGGGAGATACAGA
                                        5800
  AsnTyrGlnAsnTrpTrpArgTrpGlyMetMetLeuLeuGlyMetLeuMetThrCysSer
GGAATTATCAAAACTGGTGGAGATGGGGCATGATGCTCCTTGGGATGTTGATGACCTGTA

  IleAlaGluAspLeuTrpValThrValTyrTyrGlyValProValTrpLysGluAlaThr
GTATTGCAGAAGATTTGTGGGTTACAGTTTATTATGGGGTACCTGTGTGGAAAGAAGCAA
            5900
  ThrThrLeuPheCysAlaSerAspAlaLysSerTyrGluThrGluValHisAsnIleTrp
CCACTACTCTATTTTGTGCATCAGATGCTAAATCATATGAAACAGAAGTACATAACATCT
                                                        6000
  AlaThrHisAlaCysValProThrAspProAsnProGlnGluIleGluLeuGluAsnVal
GGGCTACACATGCCTGTGTACCCACGGACCCCAACCCACAAGAAATAGAACTGGAAAATG

  ThrGluGlyPheAsnMetTrpLysAsnAsnMetValGluGlnMetHisGluAspIleIle
TCACAGAAGGGTTTAACATGTGGAAAAATAACATGGTGGAGCAGATGCATGAGGATATAA
                                        6100
```

FIG. 7F

```
SerLeuTrpAspGlnSerLeuLysProCysValLysLeuThrProLeuCysValThrLeu
TCAGTTTATGGGATCAAAGCCTAAAACCATGTGTAAAGCTAACCCCACTCTGTGTCACTT

AsnCysThrAsnValAsnGlyThrAlaValAsnGlyThrAsnAlaGlySerAsnArgThr
TAAACTGCACTAATGTGAATGGGACTGCTGTGAATGGGACTAATGCTGGGAGTAATAGGA
              6200
AsnAlaGluLeuLysMetGluIleGlyGluValLysAsnCysSerPheAsnIleThrPro
CTAATGCAGAATTGAAAATGGAAATTGGAGAAGTGAAAAACTGCTCTTTCAATATAACCC
                                                        6300
ValGlySerAspLysArgGlnGluTyrAlaThrPheTyrAsnLeuAspLeuValGlnIle
CAGTAGGAAGTGATAAAAGGCAAGAATATGCAACTTTTTATAACCTTGATCTAGTACAAA

AspAspSerAspAsnSerSerTyrArgLeuIleAsnCysAsnThrSerValIleThrGln
TAGATGATAGTGATAATAGTAGTTATAGGCTAATAAATTGTAATACCTCAGTAATTACAC
                                        6400
AlaCysProLysValThrPheAspProIleProIleHisTyrCysAlaProAlaGlyPhe
AGGCTTGTCCAAAGGTAACCTTTGATCCAATTCCCATACATTATTGTGCCCCAGCTGGTT

AlaIleLeuLysCysAsnAspLysLysPheAsnGlyThrGluIleCysLysAsnValSer
TTGCAATTCTAAAGTGTAATGATAAGAAGTTCAATGGAACGGAAATATGTAAAAATGTCA
              6500
ThrValGlnCysThrHisGlyIleLysProValValSerThrGlnLeuLeuLeuAsnGly
GTACAGTACAATGTACACATGGAATTAAGCCAGTGGTGTCAACTCAACTGCTGTTAAATG
                                                        6600
SerLeuAlaGluGluGluIleMetIleArgSerGluAsnLeuThrAspAsnThrLysAsn
GCAGTCTAGCAGAAGAAGAGATAATGATTAGATCTGAAAATCTCACAGACAATACTAAAA

IleIleValGlnLeuAsnGluThrValThrIleAsnCysThrArgProGlyAsnAsnThr
ACATAATAGTACAGCTTAATGAAACTGTAACAATTAATTGTACAAGGCCTGGAAACAATA
                                        6700
ArgArgGlyIleHisPheGlyProGlyGlnAlaLeuTyrThrThrGlyIleValGlyAsp
CAAGAAGAGGGATACATTTCGGCCCAGGGCAAGCACTCTATACAACAGGGATAGTAGGAG

IleArgArgAlaTyrCysThrIleAsnGluThrGluTrpAspLysThrLeuGlnGlnVal
ATATAAGAAGAGCATATTGTACTATTAATGAAACAGAATGGGATAAAACTTTACAACAGG
              6800
AlaValLysLeuGlySerLeuLeuAsnLysThrLysIleIlePheAsnSerSerSerGly
TAGCTGTAAAACTAGGAAGCCTTCTTAACAAAACAAAAATAATTTTTAATTCATCCTCAG
                                                        6900
GlyAspProGluIleThrThrHisSerPheAsnCysArgGlyGluPhePheTyrCysAsn
GAGGGGACCCAGAAATTACAACACACAGTTTTAATTGTAGAGGGGAATTTTTCTACTGTA

ThrSerLysLeuPheAsnSerThrTrpGlnAsnAsnGlyAlaArgLeuSerAsnSerThr
ATACATCAAAACTGTTTAATAGTACATGGCAGAATAATGGTGCAAGACTAAGTAATAGCA
                                        7000
GluSerThrGlySerIleThrLeuProCysArgIleLysGlnIleIleAsnMetTrpGln
CAGAGTCAACTGGTAGTATCACACTCCCATGCAGAATAAAACAAATTATAAATATGTGGC

LysThrGlyLysAlaMetTyrAlaProProIleAlaGlyValIleAsnCysLeuSerAsn
AGAAAACAGGAAAAGCTATGTATGCCCCTCCCATCGCAGGAGTCATCAACTGTTTATCAA
              7100
IleThrGlyLeuIleLeuThrArgAspGlyGlyAsnSerSerAspAsnSerAspAsnGlu
ATATTACAGGGCTGATATTAACAAGAGATGGTGGAAATAGTAGTGACAATAGTGACAATG
                                                        7200
```

FIG. 7G

ThrLeuArgProGlyGlyGlyAspMetArgAspAsnTrpIleSerGluLeuTyrLysTyr
AGACCTTAAGACCTGGAGGAGGAGATATGAGGGACAATTGGATAAGTGAATTATATAAAT

LysValValArgIleGluProLeuGlyValAlaProThrLysAlaLysArgArgValVal
ATAAAGTAGTAAGAATTGAACCCCTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGG
7300

GluArgGluLysArgAlaIleGlyLeuGlyAlaMetPheLeuGlyPheLeuGlyAlaAla
TGGAAAGAGAAAAAAGAGCAATAGGACTAGGAGCCATGTTCCTTGGGTTCTTGGGAGCAG

GlySerThrMetGlyAlaAlaSerLeuThrLeuThrValGlnAlaArgGlnLeuLeuSer
CAGGAAGCACGATGGGCGCAGCGTCACTAACGCTGACGGTACAGGCCAGACAGTTACTGT
7400

GlyIleValGlnGlnGlnAsnAsnLeuLeuArgAlaIleGluAlaGlnGlnHisLeuLeu
CTGGTATAGTGCAACAGCAAAACAATTTGCTGAGGGCTATAGAGGCGCAACAGCATCTGT
7500

GlnLeuThrValTrpGlyIleLysGlnLeuGlnAlaArgValLeuAlaValGluArgTyr
TGCAACTCACGGTCTGGGGCATTAAACAGCTCCAGGCAAGAGTCCTGGCTGTGGAAAGAT

LeuGlnAspGlnArgLeuLeuGlyMetTrpGlyCysSerGlyLysHisIleCysThrThr
ACCTACAGGATCAACGGCTCCTAGGAATGTGGGGTTGCTCTGGAAAACACATTTGCACCA
7600

PheValProTrpAsnSerSerTrpSerAsnArgSerLeuAspAspIleTrpAsnAsnMet
CATTTGTGCCTTGGAACTCTAGTTGGAGTAATAGATCTCTAGATGACATTTGGAATAATA

ThrTrpMetGlnTrpGluLysGluIleSerAsnTyrThrGlyIleIleTyrAsnLeuIle
TGACCTGGATGCAGTGGGAAAAAGAAATTAGCAATTACACAGGCATAATATACAACTTAA
7700

GluGluSerGlnIleGlnGlnGluLysAsnGluLysGluLeuLeuGluLeuAspLysTrp
TTGAAGAATCGCAAATCCAGCAAGAAAAGAATGAAAAGGAATTATTGGAATTGGACAAGT
7800

AlaSerLeuTrpAsnTrpPheSerIleSerLysTrpLeuTrpTyrIleArgIlePheIle
GGGCAAGTTTGTGGAATTGGTTTAGCATATCAAAATGGCTGTGGTATATAAGAATATTCA

IleValValGlyGlyLeuIleGlyLeuArgIleIlePheAlaValLeuSerLeuValAsn
TAATAGTAGTAGGAGGCTTAATAGGTTTAAGAATAATTTTTGCTGTGCTTTCTTTAGTAA
7900

ArgValArgGlnGlyTyrSerProLeuSerLeuGlnThrLeuLeuProThrProArgGly
ATAGAGTTAGGCAGGGATACTCACCTCTGTCGTTGCAGACCCTCCTCCCAACACCGAGGG

ProProAspArgProGluGlyIleGluGluGluGlyGlyGluGlnGlyArgGlyArgSer
GACCACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGCAAGGCAGAGGCAGAT
8000

IleArgLeuValAsnGlyPheSerAlaLeuIleTrpAspAspLeuArgAsnLeuCysLeu
CAATTCGATTGGTGAACGGATTCTCAGCACTTATCTGGGACGACCTGAGGAACCTGTGCC
8100

PheSerTyrHisArgLeuArgAspLeuLeuLeuIleAlaThrArgIleValGluLeuLeu
TCTTCAGTTACCACCGCTTGAGAGACTTACTCTTAATTGCAACGAGGATTGTGGAACTTC

GlyArgArgGlyTrpGluAlaLeuLysTyrLeuTrpAsnLeuLeuGlnTyrTrpGlyGln
TGGGACGCAGGGGGTGGGAAGCCCTCAAATATCTGTGGAATCTCCTGCAATATTGGGGTC
8200

FIG. 7H

GluLeuLysAsnSerAlaIleSerLeuLeuAsnThrThrAlaIleAlaValAlaGluCys
AGGAACTGAAGAATAGTGCTATTAGCTTGCTTAATACCACAGCAATAGCAGTAGCTGAAT

ThrAspArgValIleGluIleGlyGlnArgPheGlyArgAlaIleLeuHisIleProArg
GCACAGATAGGGTTATAGAAATAGGACAAAGATTTGGTAGAGCTATTCTCCACATACCTA
8300

F
EW
MetGlyGlyLysTrpSerLys
ArgIleArgGlnGlyPheGluArgAlaLeuLeu
GAAGAATTAGACAGGGCTTCGAAAGGGCTTTGCTATAACATGGGTGGCAAGTGGTCAAAA
8400

SerSerIleValGlyTrpProLysIleArgGluArgIleArgArgThrProProThrGlu
AGTAGCATAGTAGGATGGCCTAAGATTAGGGAAAGAATAAGACGAACTCCCCCAACAGAA

ThrGlyValGlyAlaValSerGlnAspAlaValSerGlnAspLeuAspLysCysGlyAla
ACAGGAGTAGGAGCAGTATCTCAAGATGCAGTATCTCAAGATTTAGATAAATGTGGAGCA
8500

AlaAlaSerSerSerProAlaAlaAsnAsnAlaSerCysGluProProGluGluGluGlu
GCCGCAAGCAGCAGTCCAGCAGCTAATAATGCTAGTTGTGAACCACCAGAAGAAGAGGAG

GluValGlyPheProValArgProGlnValProLeuArgProMetThrTyrLysGlyAla
GAGGTAGGCTTTCCAGTCCGTCCTCAGGTACCTTTAAGACCAATGACTTATAAAGGAGCT
8600

U3
PheAspLeuSerHisPheLeuLysGluLysGlyGlyLeuAspGlyLeuValTrpSerPro
TTTGATCTCAGCCACTTTTTAAAAGAAAAGGGGGGACTGGATGGGTTAGTTTGGTCCCCA
8700

LysArgGlnGluIleLeuAspLeuTrpValTyrHisThrGlnGlyTyrPheProAspTrp
AAAAGACAAGAAATCCTTGATCTGTGGGTCTACCACACACAAGGCTACTTCCCTGATTGG

GlnAsnTyrThrProGlyProGlyIleArgPheProLeuThrPheGlyTrpCysPheLys
CAGAATTACACACCAGGGCCAGGGATTAGATTCCCACTGACCTTCGGATGGTGCTTTAAG
8800

LeuValProMetSerProGluGluValGluGluAlaAsnGluGlyGluAsnAsnCysLeu
TTAGTACCAATGAGTCCAGAGGAAGTAGAGGAGGCCAATGAAGGAGAGAACAACTGTCTG

LeuHisProIleSerGlnHisGlyMetGluAspAlaGluArgGluValLeuLysTrpLys
TTACACCCTATTAGCCAACATGGAATGGAGGACGCAGAAAGAGAAGTGCTAAAATGGAAG
8900

PheAspSerSerLeuAlaLeuArgHisArgAlaArgGluGlnHisProGluTyrTyrLys
TTTGACAGCAGCCTAGCACTAAGACACAGAGCCAGAGAACAACATCCGGAGTACTACAAA
F
9000

AspCys
GACTGCTGACACAGAAGTTGCTGACAGGGGACTTTCCGCTGGGGACTTTCCAGGGGAGGC

GTAACTTGGGCGGGACCGGGGAGTGGCTAACCCTCAGATGCTGCATATAAGCAGCTGCTT
U3 R
9100

TTCGCCTGTACTGGGTCTCTCTTGTTAGACCAGGTCGAGCCCGGGAGCTCTCTGGCTAGC

AAGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCCTCAA
9200

FIG. 7I

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/076,370	02/19/2002	Marc Alizon	2356-0011-10	2811

22852 7590 04/22/2005

FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER LLP
901 NEW YORK AVENUE, NW
WASHINGTON, DC 20001-4413

EXAMINER	
PARKIN, JEFFREY S	
ART UNIT	**PAPER NUMBER**
1648	

DATE MAILED: 04/22/2005

RECEIVED
APR 25 2005
FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP

Please find below and/or attached an Office communication concerning this application or proceeding.

Docketed 4-25-05 Attorney KJM-RMM
Case 02356.0011-10
Due Date 7-22-05 w/Ext
Action Final Resp/Appeal Due
By [signature]

/ dpm

PTO-90C (Rev. 10/03)

Office Action Summary	Application No. 10/076,370	Applicant(s) ALIZON ET AL.
	Examiner Jeffrey S. Parkin, Ph.D.	Art Unit 1648

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 03 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *09 December 2004*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *23-28 and 31-50* is/are pending in the application.

4a) Of the above claim(s) *40-50* is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *23-28 and 31-39* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

Detailed Office Action

Status of the Claims

Acknowledgement is hereby made of receipt and entry of the amendment filed 09 December, 2004. Claims 27, 28, 31, 32, and 33 were amended and new claims 40-50 introduced. Newly submitted claims 40-50 are directed to an invention that is independent or distinct from the invention originally claimed. The claims are directed toward invariant peptide sequences whereas the originally presented claims are directed toward peptidic variants. Each of the identified groups contains structurally and functionally different polypeptides. Since applicant has received an action on the merits for the originally presented invention, this invention has been constructively elected by original presentation for prosecution on the merits. Accordingly, claims 40-50 are withdrawn from further consideration as being directed towards a nonelected invention (refer to 37 C.F.R. § 1.142(b) and M.P.E.P. § 821.03). Claims 23-28 and 31-39 are currently under examination.

35 U.S.C. § 112, Second Paragraph

The previous rejection of claims 27, 28, 32, and 33 under 35 U.S.C. § 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention, is hereby withdrawn in response to applicants' amendment.

35 U.S.C. § 112, First Paragraph

The following is a quotation of the first paragraph of 35 U.S.C. § 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it

is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

Claims 23-28 and 31-39 stand rejected under 35 U.S.C. § 112, first paragraph, as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention. *In re Rasmussen*, 650 F.2d 1212, 211 U.S.P.Q. 323 (C.C.P.A. 1981). *In re Wertheim*, 541 F.2d 257, 191 U.S.P.Q. 90 (C.C.P.A. 1976). Claims 23-26 and 31 are directed toward immunogenic HIV-1 Env polypeptides of 5-150 aa comprising at least one amino acid substitution at a specified position (e.g., aa 8, 9, 90, etc.). Claims 27, 28, 32, and 33 are directed toward methodologies that require these peptides. Claims 34-39 are also directed toward immunogenic HIV-1 Env polypeptides comprising one of the aforementioned substitutions and include additional limitations pertaining to the overall peptide length (e.g., 21 aa, 43 aa, 79 aa, etc.).

As previously set forth, in order to satisfy the written description requirement, a patent specification must describe the claimed invention in sufficient detail that one skilled in the art can reasonably conclude that the inventor had **possession** of the claimed invention. See, e.g., *Vas-Cath, Inc., v. Mahurkar*, 935 F.2d at 1563, 19 U.S.P.Q.2d at 1116. The issue raised in this application is whether the original application provides adequate support for the broadly claimed genus of **immunogenic polypeptide fragments** comprising **HIV-1$_{MAL}$ epitopes** of **5-150 amino acid residues** wherein **at least one amino acid residue is substituted at one of the specified positions**. An applicant shows possession of the claimed invention by describing the claimed invention with all of its limitations using such descriptive means as words, structures,

figures, diagrams, and formulas that **fully set forth the claimed invention**. *Lockwood v. American Airlines, Inc.*, 107 F.3d 1565, 1572, 41 U.S.P.Q.2d 1961, 1966 (Fed. Cir. 1997). The claimed invention as a whole may not be adequately described where an invention is described solely in terms of a method of its making coupled with its function and there is no described or art-recognized correlation or relationship between the structure of the invention and its function. **A biomolecule sequence (e.g., epitope) described only by functional characteristic, without any known or disclosed correlation between that function and the structure of the sequence, normally is not a sufficient identifying characteristic for written description purposes, even when accompanied by a method of obtaining the biomolecule of interest.** *In re Bell*, 991 F.2d 781, 26 U.S.P.Q.2d 1529 (Fed. Cir. 1993). *In re Deuel*, 51 F.3d 1552, 34 U.S.P.Q.2d 1210 (Fed. Cir. 1995). A lack of adequate written description issue also arises if the knowledge and level of skill in the art would not permit one skilled in the art to immediately envisage the product claimed from the disclosed process. See, e.g., *Fujikawa v. Wattanasin*, 93 F.3d 1559, 1571, 39 U.S.P.Q.2d 1895, 1905 (Fed. Cir. 1995). The court noted in this decision that **a laundry list disclosure of every possible moiety does not constitute a written description of every species in a genus because it would not reasonably lead those skilled in the art to any particular species.**

An applicant may show possession of an invention by disclosure of drawings or **structural chemical formulas** that are sufficiently detailed to show that applicant was in possession of the claimed invention as a whole. An applicant may also show that an invention is complete by disclosure of sufficiently detailed, relevant identifying characteristics which provide evidence that applicant was in possession of the claimed invention, i.e., complete or

partial structure, other physical and/or chemical properties, functional characteristics when coupled with a known or disclosed correlation between function and structure, or some combination of such characteristics. For some biomolecules, examples of identifying characteristics include a nucleotide or **amino acid sequence**, chemical structure, **binding affinity**, **binding specificity**, and molecular weight. The written description requirement may be satisfied through disclosure of function and minimal structure when there is a well-established correlation between structure and function. Without such a correlation, the capability to recognize or understand the structure from the mere recitation of function and minimal structure is highly unlikely. In the latter case, **disclosure of function alone is little more than a wish for possession; it does not satisfy the written description requirement**. *Regents of the University of California v. Eli Lilly*, 119 F.3d 1559, 1566, 43 U.S.P.Q.2d 1398, 1404, 1406 (Fed. Cir. 1997), *cert. denied*, 523 U.S. 1089 (1998). *In re Wilder*, 736 F.2d 1516, 1521, 222 U.S.P.Q. 369, 372-3 (Fed. Cir. 1984). Factors to be considered in determining whether there is sufficient evidence of possession include the level of skill and knowledge in the art, partial structure, physical and/or chemical properties, functional characteristics alone or coupled with a known or disclosed correlation between structure and function, and the method of making the claimed invention.

As previously noted, perusal of the disclosure reveals the cloning and characterization of a novel human immunodeficiency virus type 1 originally designated lymphadenopathy associated virus (LAV) MAL, or LAV_{MAL}. A proviral molecular clone was obtained and complete nucleotide isolate of this sequence ascertained (see Figs. 7A-7I). The deduced amino acid sequences of the various viral structural and non-structural genes were also set forth in Figure 7. Specific envelope polypeptide fragments were set forth on p. 36

of the specification (e.g., 1-530, 34-530, 531-877, 680-700, 37-130, 211-289, 488-530, and 490-620). It should be noted that these designations actually referenced LAV_{BRU} amino acid sequences, not specific LAV_{MAL} polypeptides. Thus, the skilled artisan might conclude that applicants contemplated making and using these specific envelope polypeptides. However, the skilled artisan would not reasonably conclude that applicants were in possession of the claimed invention.

First, **the disclosure fails to identify specific HIV-1_{MAL} immunogenic fragments of the claimed lengths and substitutions.** The specification only sets forth the deduced amino acid sequences of the full-length non-structural and structural genes as set forth in Figure 7 and the specific Env fragments set forth on p. 36. Figure 3 also fails to identify immunogenic MAL peptides. This figure simply provides an amino acid comparison between MAL, BRU, ARV-2, and ELI to assess their genetic relatedness. The figure does not identify or lead the skilled artisan to any particular immunogenic fragment, particularly one carrying amino acid substitutions. Second, **the disclosure fails to perform any type of comparison wherein specific immunogenic fragments from isolate MAL are identified and acceptable amino acid substitutions are performed.** It is well-known in the art that subtle perturbations in an amino acid sequence can profoundly affect both the immunogenic and antigenic properties of any given polypeptide. Thus, the skilled artisan can only hazard a guess as to which substituted MAL fragments will remain immunogenic. Third, **the disclosure fails to provide adequate support for MAL-specific polypeptides the recited lengths (e.g., 21, 43, 79, 94, and 131 aa).** The only numerical limitations set forth in the disclosure recite immunogenic polypeptides or fusion proteins which may contain between 5 and 150 amino acids (see p. 28). Thus, support does not exist for the current size limitations. Nothing in the

disclosure directs the skilled artisan toward any particular MAL immunogenic fragment or any fragment carrying amino acid substitutions. The disclosure fails to identify those molecular determinants modulating the immunogenicity of any given polypeptide fragment. Clearly, the claimed invention simply represents an attempt by applicants to capture subject matter which was neither described nor contemplated at the time of filing. Accordingly, the skilled artisan would reasonably conclude that applicants were not in possession of the claimed invention at the time of filing.

Response to Arguments

Applicants again traverse and submit that support for the claimed substitutions can be found in Figure 3. Applicants further reference pages 16 and 23 in support. The examiner still does not concur with this assessment. Figure 3 provides the amino acid sequence alignment of four different HIV-1 isolates (e.g., BRU, ARV-2, MAL, and ELI). The purpose of this figure is to simply illustrate that while these are all HIV-1 isolates, nevertheless, they display considerable genetic diversity in the envelope region. Pages 16 and 23 also fail to provide support for the claimed invention. None of these pages sets forth any particular peptidic variants as currently claimed.

Applicants further assert that consideration of *Enzo Biochem, Inc.* v. *Gen-Probe, Inc.*, 323 F.3d 956 (Fed. Cir. 2002), is warranted. Applicants argue that the fact pattern is similar in both situations. It was noted that sufficient structural and functional information was provided in the specification. The examiner does not concur with this assessment. The problem with the current claims is that the skilled artisan would reasonably conclude that applicants were not in **possession** of the claimed invention. Nothing in the disclosure leads the skilled artisan to a polypeptide with one or more of the recited amino acid

substitutions. A more relevant piece of case law might be University of Rochester v. G.D. Searle & Co., 69 U.S.P.Q.2d 1886 (C.A.F.C. 2004). Here the court in discussing *In re Ruschig*, 379 F.2d 990 [154 U.S.P.Q. 118] (C.C.P.A. 1967), noted that although a particular compound fell within the scope of the originally filed claim, nevertheless, "it was never named or otherwise exemplified in the appellants original patent application." These facts are similar to those in the instant application wherein the applicants have failed to clearly identify or set forth any of the specific peptidic variants currently being claimed. As the Supreme court has previously cautioned, "a patent is not a hunting license. It is not a reward for the search, but compensation for its successful conclusion." *Brenner* v. *Manson*, 383 U.S. 519, 536 [148 U.S.P.Q. 689] (1966). Accordingly, the rejection is proper and hereby maintained.

Finality of Office Action

THIS ACTION IS MADE FINAL. Applicant is reminded of the extension of time policy as set forth in 37 C.F.R. § 1.136(a). **A SHORTENED STATUTORY PERIOD FOR RESPONSE TO THIS FINAL ACTION IS SET TO EXPIRE THREE MONTHS FROM THE DATE OF THIS ACTION. IN THE EVENT A FIRST RESPONSE IS FILED WITHIN TWO MONTHS OF THE MAILING DATE OF THIS FINAL ACTION AND THE ADVISORY ACTION IS NOT MAILED UNTIL AFTER THE END OF THE THREE-MONTH SHORTENED STATUTORY PERIOD, THEN THE SHORTENED STATUTORY PERIOD WILL EXPIRE ON THE DATE THE ADVISORY ACTION IS MAILED, AND ANY EXTENSION FEE PURSUANT TO 37 C.F.R. § 1.136(a) WILL BE CALCULATED FROM THE MAILING DATE OF THE ADVISORY ACTION. IN NO EVENT WILL THE STATUTORY PERIOD FOR RESPONSE EXPIRE LATER THAN SIX MONTHS FROM THE DATE OF THIS FINAL ACTION.**

Correspondence

Any inquiry concerning this communication should be directed to Jeffrey S. Parkin, Ph.D., whose telephone number is (571) 272-0908. The examiner can normally be reached Monday through Thursday from 10:30 AM to 9:00 PM. A message may be left on the examiner's voice mail service. If attempts to reach the examiner are unsuccessful, the examiner's supervisor, James C. Housel, can be reached at (571) 272-0902. Direct general status inquiries to the Technology Center 1600 receptionist at (571) 272-1600. Formal communications may be submitted through the official facsimile number which is (703) 872-9306. Hand-carried formal communications should be directed toward the customer window located in Crystal Plaza Two, 2011 South Clark Place, Arlington, VA. Applicants are directed toward the O.G. Notice for further guidance. 1280 O.G. 681. Informal communications may be submitted to the Examiner's RightFAX account at (571) 273-0908.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Respectfully,

Jeffrey S. Parkin, Ph.D.
Primary Examiner
Art Unit 1648

15 April, 2005